A Design Construct of Developing Approaches to Measure Mental Health Conditions

Md Rafiqul Islam
Department of Computer Science & Engineering,
Islamic University of Technology (IUT),
Bangladesh

Shah Jahan Miah
Information System Discipline
College of Business,
Footscray Park Campus
Victoria University
shah.miah@vu.edu.au

Abu Raihan M. Kamal
Department of Computer Science & Engineering,
Islamic University of Technology (IUT),
Bangladesh

Oliver Burmeister
School of Computing and Mathematics
Charles Sturt University

Abstract

Mental health is an important determinant of communities' well-being, influenced not only by individual attributes, but also by social and organisational environments in which people work and live. Despite studies examining mental health status among specific populations, few attempts are evident that focus on solution designs for detecting and measuring impact of mental health conditions. In this study, we develop a construct utilising design science research principles for outlining common vocabulary around the problem, and solution design relevant to a mental health management system. For the case of IT professionals, the developed construct is informed through a social-media based dataset containing more than 65,000 cells and 100 attributes potentially identifying influencing factors. Machine learning techniques are applied to the dataset to discover new findings for this specific group. It is anticipated that the analysis reported in this study would contribute in developing other electronic health management systems both for communities and healthcare professionals.

Keywords: healthcare IS; design research; mental disorder; diagnosis; occupational stress; ICT professionals.

1 Introduction

Professionals in many organizations often handle relatively stressful tasks within daily operations, affecting mental wellbeing. Mental health problems have an impact on professional communities directly through increased absenteeism; creating a negative influence on productivity and profits, as well as an increase in costs. Around half of individuals with a diagnosis of mental illness have moderate or severe occupational impairment, leading to social exclusion, poor self-esteem, and financial hardship (De

Choudhury et al. 2013). Standard, symptom-based treatments, in isolation, have inadequate effects on occupational outcomes that has led many to conclude that work outcomes for people with mental illness can be improved by increased involvement by workplaces. This involvement requires a systematic support that enables appropriate analysis, management, and exploring of solution strategies such as developing rehabilitation plans for individuals.

Mental health condition is one of the main causes of occupational dissatisfaction and sickness (French et al. 1982; Boyce et al. 2008) in the workplace (Bell & Klein, 2001). In workplaces where Information Technology (IT) professionals work, mental stress triggers when external demands and expectations exceed professionals' physical capacity, skills, and knowledge for comfortably managing problem solving and relevant tasks (Lands Bergiset et al., 1998; Joyce et al., 2009; Brohan et al. 2010). For the IT professionals, chronic exposure to stressful situations such as work overload, poor supervisory support and low input into decision-making links to a range of debilitating health outcomes (both physical and mental health issues), including depression, anxiety, emotional exhaustion, immune deficiency disorders and cardiovascular disease (Noblet & La Montagne, 2006). Although these understandings benefit social awareness and suggest treatment options, holistic solution strategies in relation to developing mental health management systems are required, and guidance for systematic solution development is still lacking.

Research in healthcare IS addresses issues regarding organizational, operational, managerial, technical and social aspects of IS in different healthcare settings. For developing systematic solutions and design knowledge much healthcare IS research takes a positivist stance (Lin et al., 2017). Design science research approaches promote views, methods, guidelines, approaches, and experiences for developing innovative IS solution artefacts such as constructs (in terms of constituting the language to specify problems and solutions vocabulary), models (in terms of abstractions for representing solutions), methods (in terms of algorithms of solutions), and instantiations (in terms of system prototypes) (Hevner et al. 2004). In this study, to address the issue of discovering factors affecting mental health conditions we outline here a construct for detecting a particular mental health condition (e.g. depression based on age group, sex, occupation, and country of the worker) among IT professionals. The proposed construct also informs a solution language by utilizing various machine learning approaches that provide an innovative concept in designing more formal systematic mental health management solutions.

Stressful working conditions impact indirectly on professionals' well-being by limiting an individual's ability to make positive changes to their lifestyle behaviours (e.g. smoking, sedentary behaviour) or by directly contributing to their negative health behaviours (Michie and Williams, 2003). At a broader scale, occupational mental issues of workforce could result in poor efficiency and throughput in the economies of countries, even impacting wider issues such as poverty and other social calamities (Cox et al. 2004; Danziger, 2009; Laitinen-Krispijn, 2000). IT professionals in particular can experience considerable occupational stress due to real time demands, equipment malfunctions, intrusion and other security threats system, system cutover problems and unanticipated disasters (Al-Saggaf et al. 2017). As IT professionals work in all sectors and in all countries, they provide a good population around which to develop a general construct.

In considering mental health of IT professionals, the following research issues are important for developing both generic problem significance and provisions for solution strategies regarding a mental health management system:

1) What exogenous and endogenous workplace factors affect mental health among IT professionals?
2) What are the attributes of IT professionals who are more vulnerable to mental health issues?
3) What are the common mental health issues diagnosed in the IT professionals and how prevalent is mental disorder?
4) What is the relationship between mental health issues and attitudes toward mental health based on cultural norms (geographic location)?
5) What is the most influential attribute contributing toward diagnosis of mental health conditions?
6) Why are people with severe mental illness excluded from work? And does it matter if people with severe mental illness work?

Understanding such concerns are prerequisite to selecting technologies for decision support solution frameworks that may better handle mental health issues.

Social media sites are abundant, offering common platforms where people express and share their personal issues using comments, photos, icons and other content. Social media data are imbued with useful information about personal well-being and life-style and many social media or organization media sites provide publicly available large datasets that contain relevant attributes of professionals. This provides a source of big data to address our questions.

Our aim in this study is to provide understanding around mental health problems and possible data-driven decision support solution approaches developed through the use of machine learning (ML) techniques. This has been shown to be effective in other areas of mental health research (Teipel et al., 2016). In DSR, research in relation to designing constructs has a significant impact on the way in which tasks and problems are comprehended for fully functional solution design (Boland, 2002; Hevner et al. 2004). Problem representation has a profound impact on design work, and theoretical constructs shape the development of models or representations of the problem domain.

The remainder of the paper is organized as follows: Section 2 presents background to the study while section 3 describes methodological details. Section 4 reports on the proposed design construct that informs mental health issue detection and guides solution design, followed by a discussion and evaluation in section 5. Finally, section 6 provides a conclusion and future directions of the study.

2 Related Works

Today's social media technologies allow people to express their views freely online which provides rich and valuable content that can enable opportunities for new insights or support information (Mislove et al., 2007). As a result, a growing number of researchers are using social media data for a range of online health, medical, and psychosocial studies (De Choudhury et al. 2013). However, it is important to carefully consider the online content of target population

for the assessment of generalisability (Choi et al. 2017). Joyce et al. (2016) identified that depression is a most common factor of mental disorders in most developed countries while carrying out a systematic meta-review examining the effectiveness of the interventions of workplace mental health. The study by Joyce et al. (2016) concluded that there are empirically supported interventions that workplaces can utilize systematic solutions to aid in the prevention of common mental illness.

Cousins et al. (2004) described the development of health and safety executive (HSE) new stress Management Standards—that offer organizations' continuous improvement through a three-phase stress preventative process—and the development of a supporting '*Indicator Tool*' (a two-phase questionnaire to assess employee perceptions of working conditions). They mentioned the HSE's aim to provide an associated methodology that will enable organizations to effectively tackle work-related stress, and subsequently reduce both its incidence and prevalence. (LaMontagne et al. 2007) discussed how job related stress has a wide range of adverse effects on mental, physical, and organizational health. These studies also emphasised how developing systems approaches for reducing the adverse impact of job stress, prevalent practice is dominated by the worker. These studies imply a clear need for designing computer-based solutions for measuring and detecting mental health conditions.

Sharma et al. (2009) developed the Global Mental Health Assessment Tool – Primary Care (GMHAT/PC) to assist general practitioners and other health professionals to make a quick, convenient, and comprehensive standardized mental health assessment. They mentioned that it may help for detecting and managing mental disorders in primary care and general health settings more effectively. Gabriel and Liimatainen (2000) represented the importance of mental health of the workplace in enhancing the economic and social integration of people with mental health difficulties. They analysed the scope and impact of mental health problems in the labour markets in five countries including Finland, Germany, Poland, United Kingdom, and United States and their work to advance and to advocate for employment for people with mental health difficulties. Cox and Cox (1993) considered the traditional approach to occupational health which points up the importance of psychosocial and organizational issues. They focused the development of a new three factors framework of concern, that is, organizational health based on health, work and organization. They addressed the organizational health problems and managed effectively within the framework that proved successful for the more tangible workplace hazards.

Although some of these studies addressed mental health issues at the workplace in general, none have considered mental stress in IT workplaces, nor have any had a focus to develop a solution design understanding, for instance, by applying data-driven methods for the purpose of mental stress detection. This paper collects social media/open sourced data to examine various attributes of tech workers and to diagnose their mental health conditions by applying ML techniques, ultimately to suggest a solution design approach for better management of mental health issues.

3 Design Methodology

The research adopted a design science research (DSR) methodology for designing the proposed artefact. DSR has received increasing attention by information systems (IS) design researchers in recent years and can produce distinct types of artefacts: constructs, models, methods, instantiations and theories (Gregor & Hevner, 2013). Baskerville et al. (2015)

described design science as representing (a) a design-science research project, (b) an artefact "build and evaluate" project such that a research project may entail, (c) the production of new knowledge from design-and-development, and (d) the creation of reports or articles describing this design-science research project. Our study follows this model to produce a new construct in the form of an artefact design and its evaluation. Baskerville et al. (2015) also suggested that DSR is an approach that enables structural guidance for researchers to create modernized or innovative system artefacts: in our study, a design construct forming a language of problems and solutions that can be applied for a real solution development.

3.1 DSR framework

The study adopted a DSR framework for the design construct comprising problem definitions and in principle solution strategies, informed by ML techniques intended to address the decision making around mental health conditions for their appropriate management. Hevner et al.'s (2004) seven guidelines provide useful criteria for defining a DSR study problem space, specifying a design based solution artefact, implementing the design solution, evaluating the design artefact and communicating study details and results. In addition, our description here is practically guided by Gregor and Hevner (2013), explicating the level of contribution to artefact abstraction and knowledge for the production of a publication (e.g. artefact design through conceptualization of ensemble artefact view defined in Miah & Gammack, 2014).

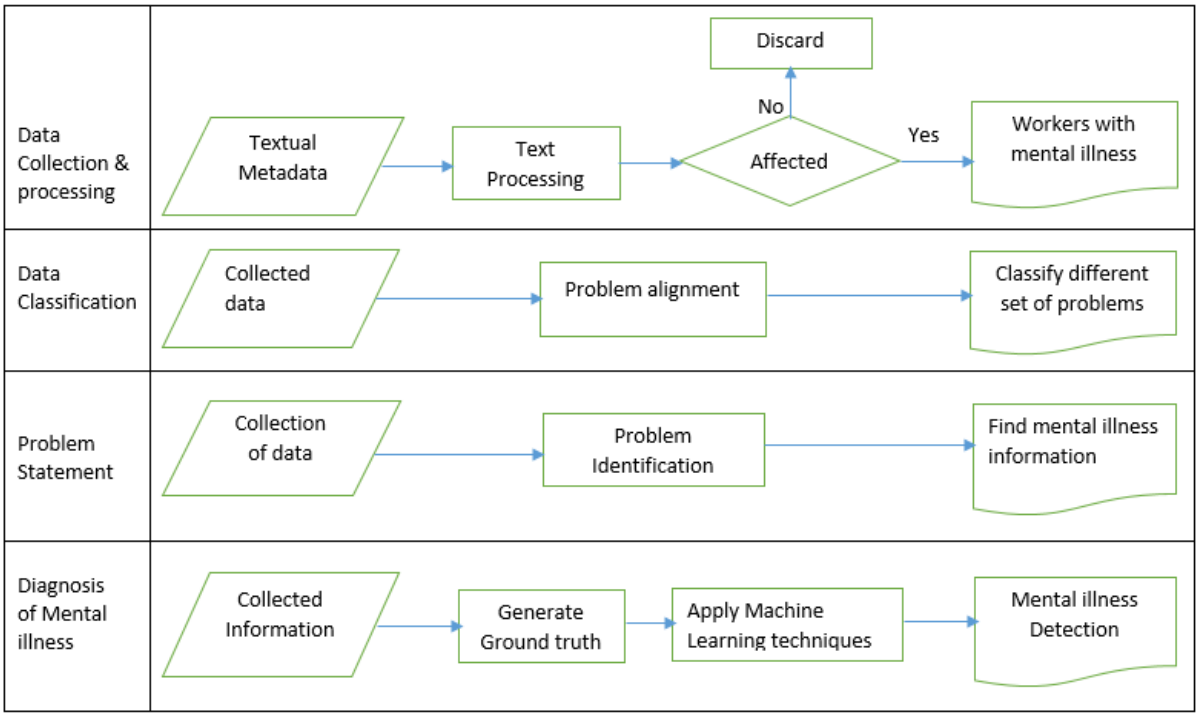

Figure 1: Our methodological steps for identifying mental health issues

We utilise four design phases namely data collection and processing, data classification, problem statement and diagnosis of mental health issues (illustrated in figure 1). This complements the approach of Miah, Vu, Gammack and McGrath (2017) in which three phases (problem identifications, artefact development and communication about the artefact) were recreated from Hevner's et al. (2004) DSR framework. We mapped these design guidelines as follows: *guideline 1* defining the type of artefact as construct, *guideline 2* is to signify the mental health problems, *guideline 3* conducting an experiment to demonstrate utility of the

technology's provisions, *guideline 4* identifying requirements of decision support for mental health issues management, *guideline 5* comparing and contrasting on the selection of technologies, *guideline 6* selecting evaluation techniques and *guideline 7* communicating the studies among academic and industry audiences.

3.2 Data collection

We used a publicly available dataset (*https://www.kaggle.com/osmi/mental-health-in-tech-2016*) containing attributes of workers working in IT/tech workplaces. The dataset contains the number of attributes of IT/Tech workplace workers. In this dataset, the total number of records in each part is approximately 10,000. We used part 1 to analyse the situation in IT workplace and part 2 to detect the percentage of IT workplace workers who are suffering from mental stress.

This dataset contains 100 columns where each column represents a unique piece of information about a tech worker. Although each of the individual columns provides valuable information for diagnosis of tech workplace workers mental health conditions, in this paper, we selected 23 columns because the other 77 columns are not strictly necessary for detecting mental health conditions of tech workplace workers. We divided our dataset into two parts to detect and examine the mental stress of IT workplace workers. Overall, these columns contain both mentally "sick" and mentally "sound" workers information. Part 1 contains 13 columns (including demographic indicators) and part 2 contains 10 columns based on the following questions which are shown below.

Overview of part 1

- Have you heard of or observed negative consequences for co-workers who have been open about mental health issues in your workplace?
- Do you believe your productivity is ever affected by a mental health issue?
- If yes, what percentage of your work time (time performing primary or secondary job functions) is affected by a mental health issue?
- Do you think that discussing a mental health disorder with previous employers would have negative consequences?
- Would you be willing to bring up a physical health issue with a potential employer in an interview?
- Would you bring up a mental health issue with a potential employer in an interview?
- Do you think that team members/co-workers would view you more negatively if they knew you suffered from a mental health issue?
- How do you describe your work position?
- How willing would you be to share with friends and family that you have a mental illness?
- What is your age?
- What is your gender?
- What country do you live in?

Overview of part 2

- Have you observed or experienced an unsupportive or badly handled response to a mental health issue in your current or previous workplace?
- Have your observations of how another individual who discussed a mental health disorder made you less likely to reveal a mental health issue yourself in your current workplace?
- Do you have a family history of mental illness?
- Have you had a mental health disorder in the past?
- Do you currently have a mental health disorder?
- If yes, what condition(s) have you been diagnosed with?
- If maybe, what condition(s) do you believe you have?
- Have you been diagnosed with a mental health condition by a medical professional?
- Have you ever sought treatment for a mental health issue from a mental health professional?
- If you have a mental health issue, do you feel that it interferes with your work when being treated effectively?

3.3 Ground Truth Dataset

We now discuss how we constructed our dataset with ground truth information (on whether the IT/Tech workers conditions are mentally sick). We use our dataset and divided into two sets (1) for the positive (YES) class (mentally sick) and (2) for the negative (NO) class (mentally sound). We justified the information in each set manually using three experts. We analysed approximately 10,000 records of IT workplace workers where 58% obtained YES and 42% obtained NO. Table 1 illustrates the overview of mental condition of IT/Tech workers.

Total number of IT/Tech workers	1,433
Mentally Sick workers	933
Mentally Sound	500

Table 1: Overview of mental condition of IT/Tech workers

3.4 Measuring Mental Illness

We acknowledge that mental health can be a variable condition, and the spectrum is not as binary as the classification of "sick and "sound" suggests, and moreover that "mental illness" is itself controversial in psychology since at least Szasz (1960). Nonetheless as with any fuzzily defined categorisation, particular classifications can be usefully distinguished and evidentially supported. As experts were used to determine the class level distinction, the detection has an authorised basis, so we preface our discussion of measurement by noting this design choice.

Figure 2 represents a measuring procedure for detecting mental health conditions using a ML technique. In this procedure, we collect data for finalizing the features of the dataset and then applied the ML technique for detecting the mental health indicators associated with anxiety, depression, bipolar disorder, personality and mood as the initial focus for our investigation.

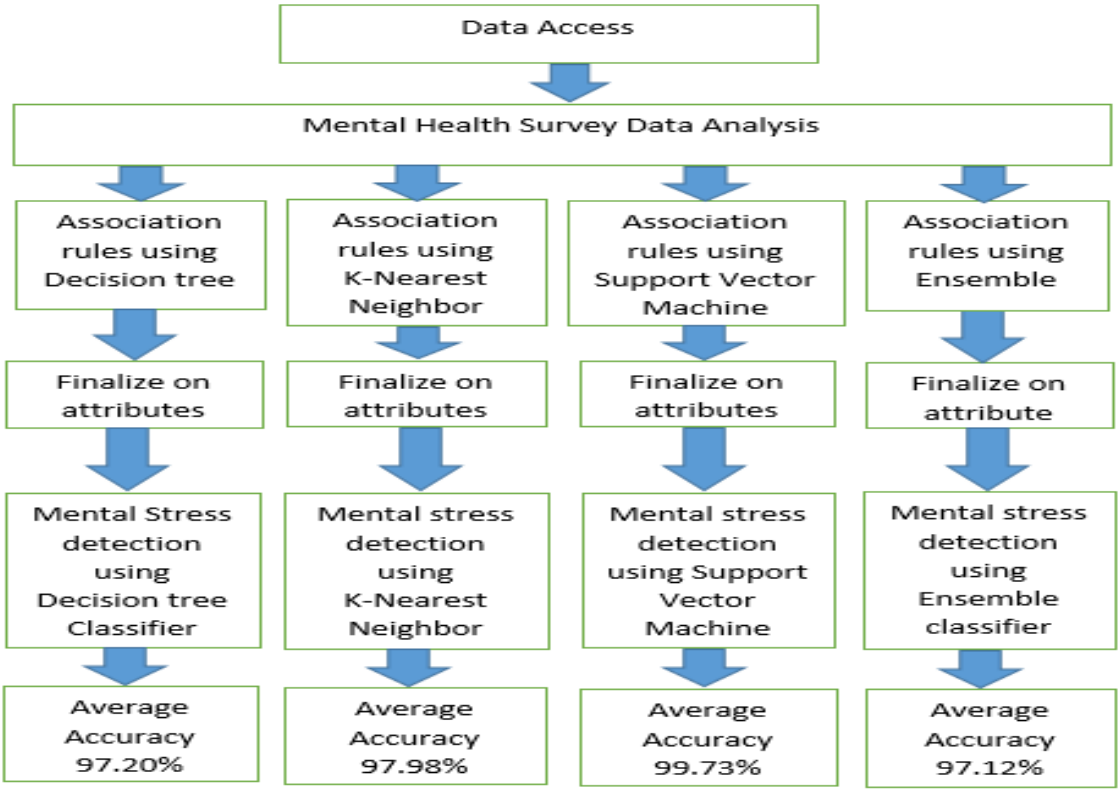

Figure 2: Detecting procedures that were used in this study

We first analyse a set of demographic attributes including age group, sex, occupation, and country that were used to represent the range of mentally sick people in IT/Tech workplace (Table 2). Attributes such as family problems, past condition, present condition, and workplace environment condition were used for detecting the mental disorder in the workplace (Table 3).

Variable	Total	Anxiety Disorder	Mood Disorder	Other Disorder
Age Group				
20-30	222	134	53	35
31-40	239	150	69	20
41-50	76	38	29	9
>50	27	13	10	4
Sex				
Male	363	207	102	54
Female	177	114	52	11
Occupation				
Back-end Developer	139	72	44	23
Supervisor/Team Lead	75	49	18	8
Executive Leadership	28	13	11	4

Variable	Total	Anxiety Disorder	Mood Disorder	Other Disorder
DevOps/SysAdmin	55	34	16	5
Dev Evangelist/Advocate	27	16	11	0
Support	26	17	6	3
One-person shop	47	20	15	12
Front-end Developer	76	55	16	5
Designer	24	16	5	3
Other	66	40	19	7
HR	3	3	0	0
Country				
United States of America	384	228	115	41
United Kingdom	62	43	17	2
Australia	16	9	6	1
Canada	23	15	3	5
Germany	10	4	5	1
Netherland	22	8	5	9
France	3	2	0	1
Denmark	2	1	0	1
Other Country	45	26	11	8

Table 2: Categorization of IT/Tech workers in various mental disorders

		Sex	Anxiety	Mood	Other	Total
Mental health disorder in the past?	YES	M	137	22	13	172
		F	82	15	3	100
	NO	M	70	14	6	90
		F	33	2	1	36
Family history of mental illness?	YES	M	86	13	8	107
		F	57	9	1	67
	NO	M	121	23	11	155
		F	58	8	3	69
Currently have a mental health disorder?	YES	M	124	22	15	161
		F	72	14	2	88
	NO	M	83	14	4	101
		F	43	3	2	48

		Sex	Anxiety	Mood	Others	Total
Mental health issue in current or previous workplace?	YES	M	107	15	9	131
		F	58	11	2	71
	NO	M	100	21	10	131
		F	57	6	2	65
Diagnosed with a mental health condition by a medical professional?	YES	M	93	16	9	118
		F	61	12	3	76
	NO	M	114	20	10	144
		F	54	5	1	60
If you have a mental health issue, do you feel that it interferes with your work when being treated effectively?	NA	M	103	20	10	133
		F	58	6	1	65
	Rarely	M	46	5	3	54
		F	21	4	1	26
	Some Times	M	47	7	5	59
		F	30	6	2	38
	Often	M	11	4	1	16
		F	6	1	0	7

Table 3: Number and proportion of IT/Tech workplace workers mental disorder into various conditions

4 Artefact Description

The proposed design construct comprises two aspects: identifying factors that affect mental health conditions at the workplace and design and evaluation of a solution approach for a particular group to showcase a solution strategy as represented in figure 3. For identifying affecting factors, we first collected open access data operating a modified API open-sourced programming code (source: https://developers.facebook.com/docs/graph-api/using-graph-api/). The collected data was in the form of textual content. We proposed an analytics methodology as illustrated in figure 4. Secondly, we used a machine learning technique for classifying the collected textual data for showcasing some findings in terms of solution design.

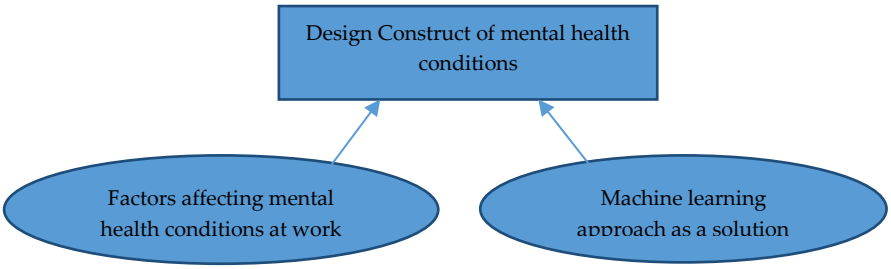

Figure 3: General structure of the proposed construct

Nunamaker et al. (1991) argued that the process of constructing solution artefacts and offering understanding of problems are important tasks for design researchers. Supporting this argument, Hevner et al. (2004) described that in design research it is imperative to provide

"the language in which problems and solutions are defined and communicated" that can inform design of "cutting-edge" IS applications.

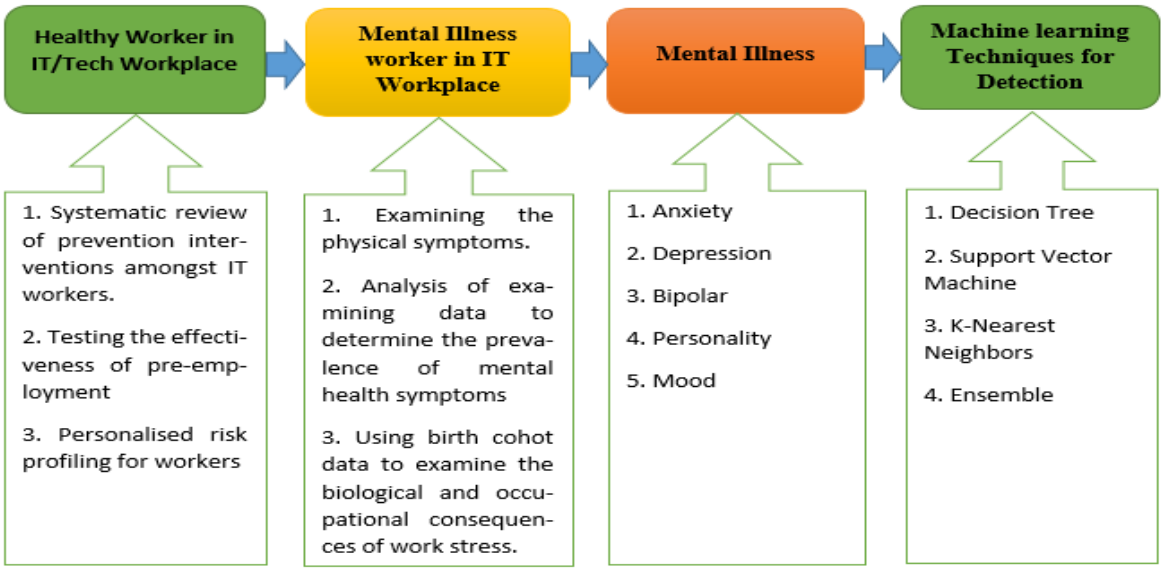

Figure 4: Method of a ML based Analytics

The proposed analytics methodology (figure 4) consists of steps such as *initial conditions of workers* (Healthy Worker in IT Workplace), assessed by a set of methods (e.g. personalized risk profiling). Next, when the workers engage their workload, their mental illness can be assessed by methods such as examining physical or mental conditions. Then the third step is to assess the mental issues, by finding the factors affecting their individual conditions. Finally the ML has been applied to validate the factors through measuring the comparison between ground truth response and ML response. The ground truth response created by three domain experts in which a binary conversion from textual data to numeric data was taken to test the ML technique's acceptability.

Machine learning in the most general conception is a set of tools and methods with the aim of modelling and understanding data. ML algorithms can learn how to perform important tasks by generalizing from examples (Oleksy, 2017). This is often feasible where manual programming is not. As more data becomes available, more ambitious problems can be tackled. These algorithms have the ability to extract information and infer structure from collections of measurements, data streams such as sound/video signal and human language (Yin et al. 2017, Burmeister et al. 2015).

Learning in general is mainly performed in two ways: knowledge acquisition and skill refinement. Thus, ML contributes to decision support systems (DSS) in two major realms: through supporting the acquisition and refinement of problem-solving knowledge and solution strategies, and through enabling DSS to adapt and evolve to be more efficient and effective in decision-making support. In this aspect we gained knowledge for investigating how ML techniques can be suitable for good decision support and useful for knowledge acquisition in their ability to adapt.

5 Discussion and Evaluation

The paper so far has presented a construct design for informing important mental health issues and possible solution provision so a fully-functional decision support application in future can be developed. We promoted an idea of using online available data to assess potential of the factors affecting mental health at IT workplace and machine learning based solution design. We discussed the performance of different ML classifiers in detecting mentally ill workers in a shorter time. The experiment was conducted using MATLAB (MATLAB is a productive software environment for researchers to analyse and visualize data with latest techniques. We used the latest version of 2016b, for better outcome). We applied four major ML classifiers such as Decision trees (it is method for using a tree-like graph or model of decision aspects), K-Nearest Neighbour (KNN), Support Vector Machine (Support Vector Machine (SVM) is a discriminative classifier, is one of the most popular machine learning algorithms), and Ensemble as mentioned earlier.

Each ML classifier technique has sub-classifiers such as Decision trees- Simple Tree, Medium Tree, and Complex Tree; SVM- Linear, Quadratic, Cubic, Fine Gaussian, Medium Gaussian, and Coarse Gaussian; KNN- Fine, Medium, Coarse, Cosine, Cubic and Weighted, Ensemble-Boosted tree, Bagged tree, Subspace discriminant, Subspace KNN, RUSBoosted tree. Performances of these classifiers have been computed by using the evaluation matrices parameters such as precision, recall, and F-measure (F-measure has a common use in data mining projects. It combines precision and recall for the harmonic mean of precision and recall).

All the evaluation metrics that we used are defined as follows.

Precision is the ratio of true positives to the cases that are predicted as positive. It is the percentage of selected cases that are correct.

$$Precision\ (P) = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

Recall is the ratio of true positives to the cases that are actually positive. It is the percentage of corrected cases that are selected.

$$Recall\ (R) = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

F-measure is the mean of Precision and Recall. It takes both false positives and false negatives into an account. F-measure is calculated as:

$$FMeasure = 2\frac{PR}{P + R}$$

The experiment was carried out by using 10-fold cross-validation on all testing data sets. For every ML classifier, we found the value of its sub-classifier which persists to precision, recall and F-measure (refer to table 4). The result has been graphically represented in figures 5 and 6.

Algorithm	*Precision*	*Recall*	*F-measure*	*Accuracy*
Complex Tree	0.994623	0.991425	0.993022	99
Medium Tree	0.994577	0.982851	0.988679	98.5
Simple Tree	0.957974	0.952840	0.955400	94.1
Fine KNN	0.997851	0.995712	0.996781	99.5
Medium KNN	0.988159	0.983922	0.986036	98.1
Coarse KNN	0.933537	0.978563	0.955520	94
Cosine KNN	0.997826	0.983922	0.990825	98.7
Cubic KNN	0.988159	0.983922	0.986036	98.1
Weighted KNN	0.997851	0.995712	0.996781	99.5
Linear SVM	0.997860	1	0.998929	99.9
Quadratic SVM	0.996794	1	0.998394	99.8
Cubic SVM	0.996794	1	0.998394	99.8
Fine Gaussian SVM	0.997851	1	0.998924	99.5
Medium Gaussian SVM	0.996788	0.998927	0.997856	99.7
Coarse Gaussian SVM	0.996791	0.998928	0.997858	99.7
Ensemble Boosted Tree	0.996784	0.996784	0.996784	99.5
Ensemble Bagged Tree	0.996791	0.998928	0.997858	99.7
Ensemble Subspace Discriminant	0.95010	0.979635	0.964643	95.3
Ensemble Subspace KNN	0.904244	0.981779	0.941418	92
Ensemble RUSBoosted Tree	0.995694	0.991425	0.993555	99.1

Table 4: Precision and recall corresponding to best F-measure of the different classifier

Figure 5 represents the result of precision and recall values of different classifiers where the x-axis represents the different sub-classifiers of Decision Tree, KNN, SVM and Ensemble techniques. We mentioned (1-20) as a symbol of Decision trees- Simple Tree, Medium Tree, and Complex Tree; SVM- Linear, Quadratic, Cubic, Fine Gaussian, Medium Gaussian, and Coarse Gaussian; KNN- Fine, Medium, Coarse, Cosine, Cubic and Weighted, Ensemble-Boosted tree, Bagged tree, Subspace discriminant, Subspace KNN, RUSBoosted tree respectively and y-axis represents the range of value.

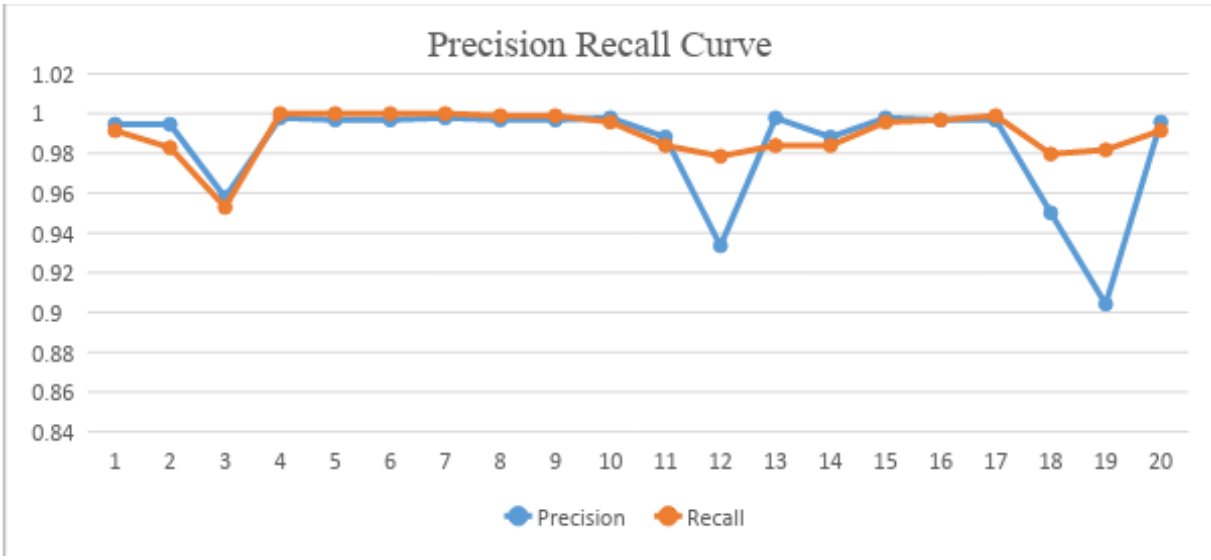

Figure 5: Precision-recall curve for the 20 sub-classifiers of machine learning techniques

Similarly, Figure 6 showed the result of F-Measure where the x-axis represents the name of different sub-classifiers and y-axis represents the percentage of the sub-classifier result.

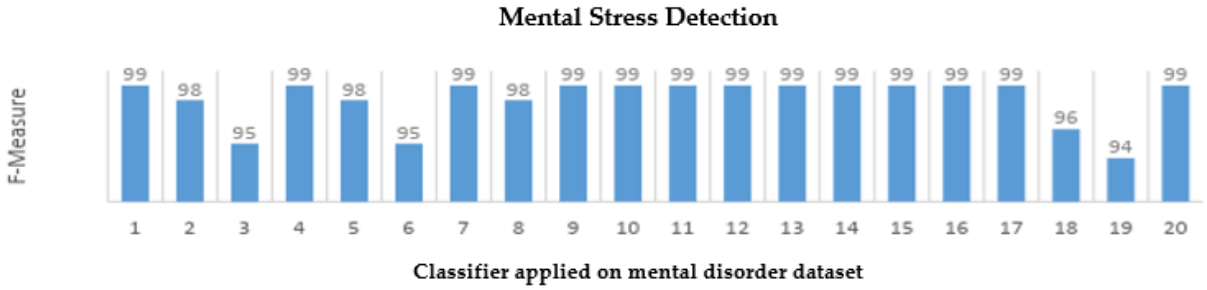

Figure 6: Performance comparisons among the classifiers

Using the above ML classification techniques, we examined detection performance of mentally ill IT workplace workers, so as to comprehend the significance of different feature types. We already demonstrated the results of various characterizations with various proportions of four features. The outcome showed that, in our test set, the best performing model is SVM. In addition, for all of the features precision, recall and F-measure calculation, SVM gives the most outstanding outcome relating to the class of mental disorder indicative of IT workplace workers.

Moreover, in figures 7 and 8, we have shown the various range of mental disorders values based on age group and sex, where x-axis represents the various names of the mental disorders and y-axis represents the number of people who are feeling anxious and moody.

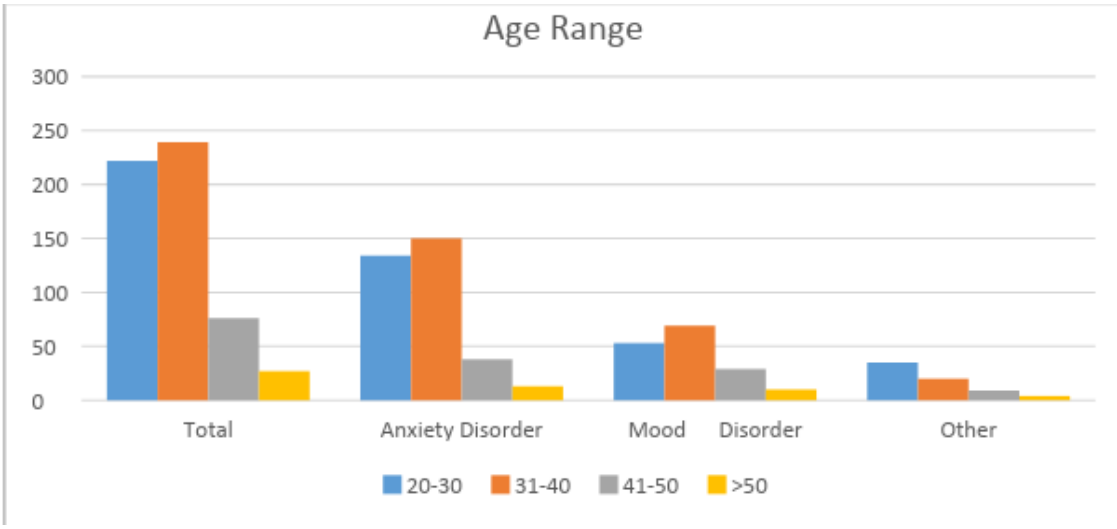

Figure 7: The range of age group in IT/Tech workers with mental disorders

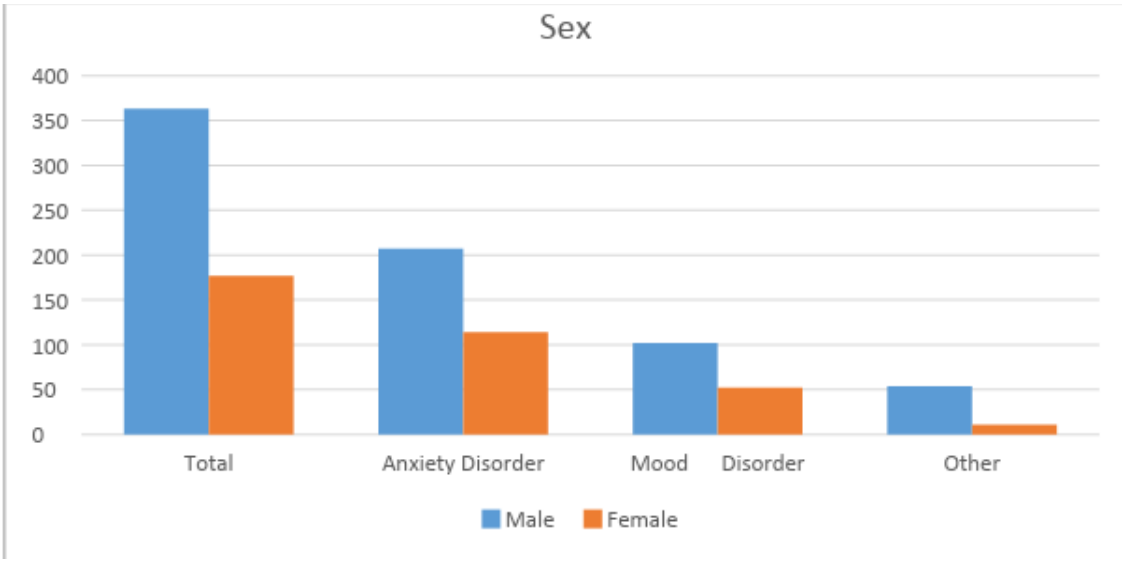

Figure 8: The number of male and female in IT/Tech workers with mental disorders

Similarly, figures 9 and 10 demonstrate the result of the range of mental disorder values based on occupation/positions such as back-end developer, supervisor/team leader, executive, system admin, dev evangelist, one person shop officer, font-end developer, designer, and HR, and country such as USA, UK, Australia, Canada, Germany, Netherland, and France, where the x-axis represents the designation of IT workplace workers and their country respectively and the y-axis represents the number of people who are feeling anxiety, mood etc. In figure 9, we examined all of the occupational positions in the IT workplace for calculating the mental disorders. It was observed that for all of the mental disorders, back-end developers are suffering the most. We trust that our present study has laid the ground for future research on deductions and revelation of new data in view of cause-event correlation. A limit in the data set used is the small number of occupation classifications. The IT workplace is much more diverse than these few classifiers suggest, and they vary between countries (Al-Saggaf et al., 2017).

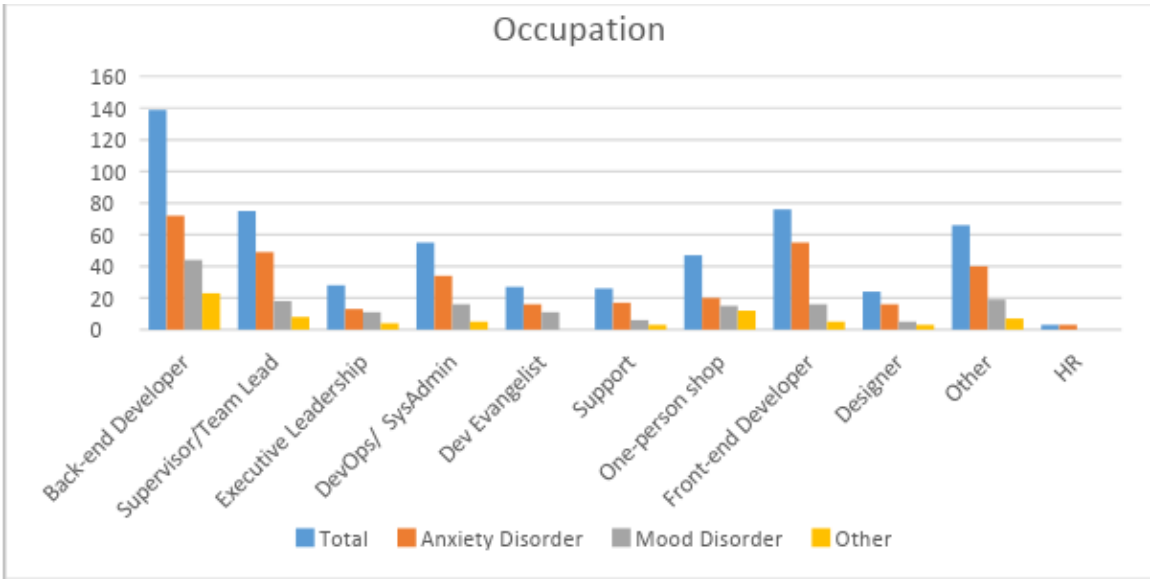

Figure 9: The number of people in the different sections of the IT/Tech workplace with mental disorders.

Figure 10 shows the populations of some countries who are suffering from different mental disorders in the IT/Tech workplace. The country with the largest population (67%) is the USA with anxiety, mood and other disorders, whereas Denmark has the smallest population. Apart from the USA, the largest countries are UK, Australia, and Canada respectively. However, interpretation of the figure can be misleading, because it is population based and based upon the limited sample in the publicly available data set. For instance, anxiety and depression are the two most prevalent mental illnesses experienced in Australia (Birch, 2015; Sayers et al., 2017). Similarly, in figure 8 anxiety is the most prevalent in IT workers. Mood in this data set includes, but is not limited to, depression. A more representative data set would therefore be likely to show different results to what is depicted in figure 10.

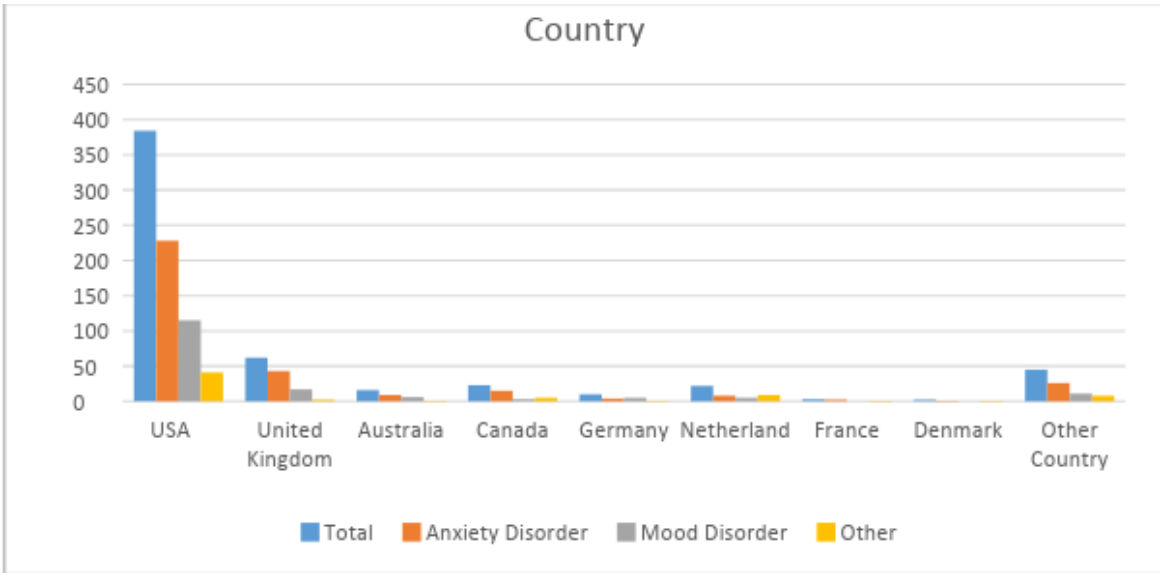

Figure 10: The number of people in the different countries of the IT/Tech workplace with mental disorders

6 Conclusion

The paper represents basic requirements of developing a systematic IS solution for establishing common vocabulary of problems and solution technique selection. In our study we used a representative case of IT professionals. For promoting an understanding on mental health issues and possible technology selection for decision support solution design, we have exhibited the capability of using IT/Tech workplace survey data for measuring and detecting mental disorders among tech workers. To give a clear understanding of our work, the research challenges were stated at the start of this paper. The proposed analytics performed on the selected dataset, provide some insight on the research challenges. Below is the summary of our findings:

- What exogenous and endogenous workplace factors affect the mental health of tech worker?

It was revealed that there are four key factors that significantly contribute toward the mental health condition of tech workers. These factors include 'minimal risk', 'workplace concern', 'activity concern', and 'emotional stresses.

- What are the attributes of tech workers who are more vulnerable to mental health issues?

Although, most of the workers working in technical departments are prone to mental health issues, our study found that the worker with 'family history of mental stress', 'past condition of mental stress', are more likely to develop mental health issues. In addition, the current workplace environment can act as a catalyst to further magnify the chances of mental health of already vulnerable workers.

- What are the common mental health issues diagnosed in the tech workers and how many mental disorder episodes they had experienced?

Proposed analytics performed on the selected dataset suggest various mental health issues that tech workers could possibly be presented with. Common mental health issues include 'anxiety disorder' (generalized, social, phobia, etc.), 'mood disorder' ('depression', 'bipolar disorder', etc.) and other kinds of disorders such as ('personality disorder', 'seasonal affective disorder', 'burn out', 'PDD-NOS', 'combination of physical impairment' (strongly near-sighted) with a possibly mental one (MCD / "ADHD"), 'substance use disorder' , 'stress response syndromes', 'sexual addiction', 'obsessive-compulsive disorder', and 'post-traumatic stress'). These results showing a relationship between mental illness and physical impairment are recognised more widely, as seen in the Equally Well initiative recently launched by the Australian National Mental Health Commission to better address the physical needs of people experiencing mental illness (Roberts et al. 2018).

Analytics presented in Table 2 above suggested that workers in the age bracket of 20-30 years are more likely to get anxiety. This finding is supported by an OECD report (OECD, 2012) on the relationship between work and mental health, which reported that mental health workplace challenges tend to arise at an early age. This could possibly due to lack of their experience of dealing with high-stress tech environment and other pressures at the early stages of their professional career. The situation could be better with the age as workers get more experience and find ways to deal with work related anxiety and depression.

- What is the relationship between mental health issues and attitudes toward mental health based on the geographic location in the global workplace?

The relationship of mental health issues with geographic locations varies country to country. Table 3 presents information on mental health condition of workers working in various countries.

- What is the most influential attribute contributing toward diagnosis of mental health condition in the tech workplace?

Although we considered a total of 23 attributes for diagnosis of mental health conditions of tech workers, the most significant attributes are found to be family history of mental illness, personal history of mental illness, present situation of mental health and the working environment. We trust that these attributes provide valuable information on the influential attributes contributing toward the diagnosis of mental health condition in the tech workplace. In this paper, we make an attempt to get an insight of mental health condition of tech workers by applying various ML techniques. Our findings suggest that all the classifiers results are almost 99%.

In future work, we will extend the literature in the area of growing attention called 'Mental Health Mining', in particular in health related knowledge discovery. The suitability of mental illness data for disease surveillance will be demonstrated along with its potential limitations in a case study of mental illness around the world. Additionally, the state-of-art and empirically validated machine learning algorithm will be employed in order to gain insight into the most common mental illnesses around the world. We will use further techniques to extract paraphrases from more types of mental disorder features. Also, we will be using more datasets to verify the efficiency and effectiveness of ML techniques to subsequently improve the predictive accuracy. This will prove the way for a larger scale investigation into the topic in the upcoming research. As a result, the benefits of high quality input data as well as the innovative visualization tool for patterns exploration will considerably facilitate knowledge discovery. Finally, the literature on social media data mining for mental health applications will be extended with a new practical application to the current matters of concern, including depression and anxiety.

References

Al-Saggaf, Y., Burmeister, O. K., & Schwartz, M. (2017). Qualifications and ethics education: the views of ICT professionals. Australasian Journal of Information Systems, 21, DOI: 10.3127/ajis.v21i0.1365.

Baskerville, R. L., Kaul, M., & Storey, V. C. (2015). Genres of Inquiry in Design-Science Research: Justification and Evaluation of Knowledge Production. *MIS Quarterly*, *39*(3), 541-564.

Bell, B. S., & Klein, K. J. (2001). Effects of disability, gender, and job level on ratings of job applicants. *Rehabilitation Psychology*, *46*(3), 229.

Bender, A., & Farvolden, P. (2008). Depression and the workplace: a progress report. *Current psychiatry reports*, *10*(1), 73-79.

Birch, J. (2015). *Beyondblue* NewAccess Demonstration Independent Evaluation, Summary of Findings, Reporting to August 2015, available at https://www.beyondblue.org.au/docs/default-source/research-project-

files/bw0353_beyondblue-newaccess-demonstration-independent-evaluation.pdf?sfvrsn=7e1050ea_0 accessed on 6 June 2018

Boland, R. (2002). Design in the punctuation of management action. Managing as designing: Creating a vocabulary for management education and research, In: Boland, R. (ed.) Frontiers of Management Workshop. Weatherhead School of Management, June 14-15, 106-112.

Boyce, M., Secker, J., Johnson, R., Floyd, M., Grove, B., Schneider, J., & Slade, J. (2008). Mental health service users' experiences of returning to paid employment. *Disability & Society, 23*(1), 77-88.

Brohan, E., Henderson, C., Little, K., & Thornicroft, G. (2010). Employees with mental health problems: Survey of UK employers' knowledge, attitudes and workplace practices. *Epidemiology and Psychiatric Sciences, 19*(4), 326-332.

Burmeister, O. K., Islam, M. Z., Dayhew, M., & Crichton, M. (2015). Enhancing client welfare through better communication of private mental health data between rural service providers. *Australasian Journal of Information Systems*, 19, 1-14. doi:http://dx.doi.org/10.3127/ajis.v19i0.1206

Bzdok, D., & Meyer-Lindenberg, A. (2017). Machine learning for precision psychiatry. *arXiv preprint arXiv:1705.10553*.

Choi, I., Milne, D. N., Glozier, N., Peters, D., Harvey, S. B., & Calvo, R. A. (2017). Using different Facebook advertisements to recruit men for an online mental health study: engagement and selection bias. *Internet Interventions, 8*, 27-34.

Cousins, R., Mackay, C. J., Clarke, S. D., Kelly, C., Kelly, P. J., & McCaig, R. H. (2004). 'Management standards' work-related stress in the UK: Practical development. *Work & Stress, 18*(2), 113-136.

Cox, T., & Cox, S. (1993). Occupational health: Control and monitoring of psychosocial and organisational hazards at work. *Journal of the Royal Society of Health, 113*(4), 201-205.

Cox, T., Leka, S., Ivanov, I., & Kortum, E. (2004). Work, employment and mental health in Europe. *Work & Stress, 18*(2), 179-185.

Danziger, S., Frank, R. G., & Meara, E. (2009). Mental illness, work, and income support programs. *American Journal of Psychiatry, 166*(4), 398-404.

De Choudhury, M., Counts, S., & Horvitz, E. (2013). *Social media as a measurement tool of depression in populations.* Paper presented at the Proceedings of the 5th Annual ACM Web Science Conference.

French, J. R., Caplan, R. D., & Van Harrison, R. (1982). *The mechanisms of job stress and strain* (Vol. 7): Chichester [Sussex]; New York: J. Wiley.

Gabriel, P., & Liimatainen, M.-R. (2000). Mental health in the workplace: Introduction, executive summaries, Cornel University ILR School, DigitalCommons, Available: http://digitalcommons.ilr.cornell.edu/gladnetcollect/223/, accessed on 8 June 2018.

Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly, 37*(2).

Hatch, S. L., Harvey, S. B., Dandeker, C., Burdett, H., Greenberg, N., Fear, N. T., & Wessely, S. (2013). Life in and after the Armed Forces: social networks and mental health in the UK military. *Sociology of Health & Illness, 35*(7), 1045-1064.

Hu, H., Li, J., Wang, H., & Daggard, G. (2006). *Combined gene selection methods for microarray data analysis.* Paper presented at the International Conference on Knowledge-Based and Intelligent Information and Engineering Systems.

Hu, H., Li, J., Wang, H., Daggard, G., & Shi, M. (2006). *A maximally diversified multiple decision tree algorithm for microarray data classification.* Paper presented at the Proceedings of the 2006 workshop on Intelligent systems for bioinformatics-Volume 73.

Huang, J., Peng, M., Wang, H., Cao, J., Gao, W., & Zhang, X. (2017). A probabilistic method for emerging topic tracking in microblog stream. *World Wide Web, 20*(2), 325-350.

Joyce, S., Modini, M., Christensen, H., Mykletun, A., Bryant, R., Mitchell, P. B., & Harvey, S. B. (2016). Workplace interventions for common mental disorders: a systematic meta-review. *Psychological Medicine, 46*(4), 683-697.

Joyce, T., McMillan, M., & Hazelton, M. (2009). The workplace and nurses with a mental illness. *International Journal of Mental Health Nursing, 18*(6), 391-397.

Kabir, M. E., Wang, H., & Bertino, E. (2012). A role-involved purpose-based access control model. *Information Systems Frontiers, 14*(3), 809-822.

Khalil, F., Wang, H., & Li, J. (2007). *Integrating markov model with clustering for predicting web page accesses.* Paper presented at the Proceeding of the 13th Australasian World Wide Web Conference (AusWeb07).

Laitinen-Krispijn, S., & Bijl, R. (2000). Mental disorders and employee sickness absence: the NEMESIS study. *Social Psychiatry and Psychiatric Epidemiology, 35*(2), 71-77.

LaMontagne, A. D., Keegel, T., & Vallance, D. (2007). Protecting and promoting mental health in the workplace: developing a systems approach to job stress. *Health Promotion Journal of Australia, 18*(3), 221-228.

Landsbergis, P. A., Schnall, P. L., Deitz, D. K., Warren, K., Pickering, T. G., & Schwartz, J. E. (1998). Job strain and health behaviors: results of a prospective study. *American journal of Health Promotion, 12*(4), 237-245.

Lin, Y.-K., Chen, H., Brown, R. A., Li, S.-H., & Yang, H.-J. (2017). Healthcare predictive analytics for risk profiling in chronic care: a bayesian multitask learning approach. *MIS Quarterly, 41*(2).

Miah, S.J., Vu, QH, Gammack, J. & McGrath, M. (2017). A Big-Data Analytics Method for Tourist Behaviour Analysis, *Information and Management*, 54, 771-785

Miah, S.J. & Gammack, J. (2014). Ensemble Artifact Design for Context Sensitive Decision Support, *Australasian Journal of Information Systems* 18(2), 5-20

Michie, S., & Williams, S. (2003). Reducing work related psychological ill health and sickness absence: a systematic literature review. *Occupational and Environmental Medicine, 60*(1), 3-9.

Mislove, A., Marcon, M., Gummadi, K. P., Druschel, P., & Bhattacharjee, B. (2007). *Measurement and analysis of online social networks.* Paper presented at the Proceedings of the 7th ACM SIGCOMM conference on Internet measurement.

Noblet, A., & LaMontagne, A. D. (2006). The role of workplace health promotion in addressing job stress. *Health Promotion International, 21*(4), 346-353.

Nunamaker Jr, J. F., Chen, M., & Purdin, T. D. (1990). Systems development in information systems research. *Journal of Management Information Systems, 7*(3), 89-106.

OECD. (2012). Sick on the Job? Myths and Realities about Mental Health and Work. Paris: OECD Publishing Retrieved from http://dx.doi.org/10.1787/9789264124523-en.

Oleksy, T. A. (2017). Machine Learning methods for mood disorder decision support. The University of Bergen.

Roberts, R., Ryan, K., Burmeister, O., and Maylea, C. (2018) A National Priority: The physical health of people living with mental illness, *Australian Nursing & Midwifery Journal*, 25(10), 42.

Sayers, J. M., Cleary, M., Hunt, G. E., & Burmeister, O. K. (2017). Service and infrastructure needs to support recovery programmes for Indigenous community mental health consumers. *International Journal of Mental Health Nursing*, 26, 142–150.

Sharma, V. K., & Copeland, J. R. (2009). Detecting mental disorders in primary care. *Mental Health in Family Medicine*, 6(1), 11.

Sun, X., Wang, H., Li, J., & Zhang, Y. (2011). Injecting purpose and trust into data anonymisation. *Computers & Security, 30*(5), 332-345.

Szasz, T. S. (1960). The myth of mental illness. *American Psychologis*t, 15(2), 113-118. http://dx.doi.org/10.1037/h0046535

Teipel, S., Babiloni, C., Hoey, J., Kaye, J., Kirste, T., & Burmeister, O.K. (2016). Information and communication technology solutions for outdoor navigation in dementia. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 12(6), 695-707. doi:10.1016/j.jalz.2015.11.003

Umanandhini, M. D., & Kalpana, G. (2017). Survey on stress types using data mining algorithms. *Stress, 4*(04).

Von Alan, R. H., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly, 28*(1), 75-105.

Wang, Y., Li, H., Wang, H., Zhou, B., & Zhang, Y. (2015). *Multi-window based ensemble learning for classification of imbalanced streaming data.* Paper presented at the International Conference on Web Information Systems Engineering.

Yin, X.-X., Hadjiloucas, S., & Zhang, Y. (2017). Pattern classification *Pattern Classification of Medical Images: Computer Aided Diagnosis* (pp. 93-128): Springer.

Yin, X., Hadjiloucas, S., & Zhang, Y. (2016). Classification of THz pulse signals using two-dimensional cross-correlation feature extraction and non-linear classifiers, *Computer methods and programs in biomedicine, 127*, 64-82.