# ONLINE RATING SCALE FOR THE LIFE DISTRESS INVENTORY

**SUWANNEE SUKSAWAS**

**A THEMATIC PAPER SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE
(TECHNOLOGY OF INFORMATION SYSTEM MANAGEMENT)
FACULTY OF GRADUATE STUDIES
MAHIDOL UNIVERSITY
2011**

Thematic Paper
entitled

# ONLINE RATING SCALE FOR THE LIFE DISTRESS INVENTORY

……………………………………..
Miss. Suwannee Suksawas
Candidate

……………………………………..
Assoc.Prof. Sucheera Phattharayuttawat,
Ph.D. (Psychology)
Major advisor

……………………………………...
Lect. Songpol Ongwattanakul,
Ph.D.(Computer Engineering)
Co-advisor

……………………………………..
Prof. Banchong Mahaisavariya,
M.D.,Dip.Thai Board of Orthopedics
Dean
Faculty of Graduate Studies
Mahidol University

……………………………………..
Asst. Prof. Rawin Raviwongse,
Ph.D. (Engineering Management)
Program Director
Master of Science Program in
Technology of Information System
Management
Faculty of Engineering
Mahidol University

Thematic Paper
entitled

**ONLINE RATING SCALE FOR THE LIFE DISTRESS INVENTORY**

was submitted to the Faculty of Graduate Studies, Mahidol University
for the degree of Master of Science
(Technology of Information System Management)
on
February 9, 2011

…………………………………..
Miss. Suwannee Suksawas
Candidate

……………………………………
Asst.Prof. Warakorn Charoensuk,
Ph.D. (Electrical Engineering)
Chair

……………………………………
Assoc. Prof. Sucheera Phattharayuttawat,
Ph.D. (Psychology)
Member

……………………………………
Mr. Supanat Kitcharoensakkul,
Ph.D (Computer Science)
Member

……………………………………
Lect. Songpol Ongwattanakul,
Ph.D. (Computer Engineering)
Member

……………………………………..
Prof. Banchong Mahaisavariya,
M.D.,Dip.Thai Board of Orthopedics
Dean
Faculty of Graduate Studies
Mahidol University

……………………………………..
Asst. Prof. Rawin Raviwongse,
Ph.D. (Engineering Management)
Dean
Faculty of Engineering
Mahidol University

# ACKNOWLEDGEMENTS

The success of this thematic paper can be attributed to the extensive support and assistance from my major advisor, Assoc. Prof. Sucheera Phattharayuttawat, my co-advisor, Dr.Songpol Ongwattanakul. I deeply thank them for their valuable advice and guidance in this research.

I would like to thank all the teachers and staffs at the Technology of Information System Management program for their kind support during the time of my study.

In addition, I would like to thank every participant who cooperated with the research procedures, participated and devoted their time.

Finally, I am grateful to my family for their understanding. They give me plenty of time to accomplish this thematic paper. Moreover, they support me unconditionally with love and care. I dedicate all the credits of this work to my mother, my father, and all teachers who have taught me since my childhood.

Suwannee Suksawas

ONLINE RATING SCALE FOR THE LIFE DISTRESS INVENTORY

SUWANNEE SUKSAWAS 5036639 EGTI/M

M.Sc. (TECHONOLOGY OF INFORMATION SYSTEM MANAGEMENT)

THEMATIC PAPER ADVISORY COMMITIEE: SUCHEERA PHATTHARAYUTTAWAT, Ph.D. (PSYCHOLOGY), SONGPOL ONGWATTANAKUL, Ph.D. (COMPUTER ENGINEERING)

## ABSTRACT

This study is an extension to the development of an online rating scale for The Life Distress Inventory questionnaire which was originally paper-pencil based. The two main objectives of this study were to develop the Life Distress Inventory (LDI), computerized version, and to store the information into a database which clinicians, researchers, and specialists in psychology could use to generate reports for future use. An online rating scale for the LDI questionnaire was developed to support the data gathering process, which allows the public to access and fill in the LDI questionnaire online. Additional information, such as some personal data, is required for every questionnaire. All participants must fill in their personal information and 20 questions from the LDI questionnaire via the internet.

The analysis involved a data mining technique called clustering. Clustering is a common tool used to divide data into groups (clusters). Data with the same or similar features are arranged in the same group. Segmentation of the data can be interpreted as knowledge governed by the data. In general, this work is in line with the field of Knowledge Discovery in Database (KDD).

The results found that specialists had high satisfaction with the LDI computerized version as an online version for The Life Distress Inventory, and the other results reflect that it will be highly promising to discover hidden knowledge from the online version of The Life Distress Inventory database. The system is more knowledgeable because of an excellent database system and web application as a measure of life distress inventory, which allows for an interaction between clinician, researcher, and all participants. All data is more up-to-date when analyzed by the clinicians, researchers, and psychology specialists.

KEY WORDS : THE LIFE DISTRESS INVENTORY/
DATA MINING / CLUSTERING

88 pages

แบบสำรวจสถานการณ์ที่ก่อให้เกิดความเครียดระบบออนไลน์
ONLINE RATING SCALE FOR THE LIFE DISTRESS INVENTORY

สุวรรณี สุขสวัสดิ์ 5036639 EGTI/M

วท.ม (เทคโนโลยีการจัดการระบบสารสนเทศ)

คณะกรรมการที่ปรึกษาสารนิพนธ์: สุชีรา ภัทรายุตวรรตน์, Ph.D. (PSYCHOLOGY)
ทรงพล องค์วัฒนกุล Ph.D. (COMPUTER ENGINEERING)

บทคัดย่อ

การศึกษาวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาแบบสำรวจสถานการณ์ที่ก่อให้เกิดความเครียด (The Life Distress Inventory: LDI) รูปแบบเว็บแอบพลิเคชั่น ซึ่งแบบสอบถามฉบับดั้งเดิมนั้นอยู่ในรูปแบบกระดาษ และทำการสกัดองค์ความรู้ที่ซ่อนอยู่ในธนาคารข้อมูลของแบบสำรวจสถานการณ์ที่ก่อให้เกิดความเครียด โดยระบบธนาคารข้อมูลLDI ถูกพัฒนาขึ้นเพื่อช่วยในการเก็บข้อมูล ซึ่งบุคคลทั่วไปสามารถเข้ามาทำแบบทดสอบ LDI ฉบับคอมพิวเตอร์ได้ด้วยวิธีการออนไลน์ผ่านอินเตอร์เน็ต ผู้ทำแบบทดสอบจะต้องกรอกข้อมูลส่วนตัวและทำแบบทดสอบจำนวน 20 ข้อและทำการจัดเก็บข้อมูลที่ได้จากการตอบแบบสอบถามทั้งหมดลงในฐานข้อมูล

การวิเคราะห์ข้อมูลด้วยวีธีการจัดกลุ่มข้อมูล (clustering) เป็นหนึ่งในเทคนิคของการทำเหมืองข้อมูล (Data mining) ซึ่งการวิเคราะห์ข้อมูลด้วยวีธี Clustering นี้เป็นการทำเหมืองข้อมูลโดยจะแบ่งชุดข้อมูลออกเป็นกลุ่ม (cluster) นำข้อมูลที่มีคุณลักษณะเหมือนกัน หรือคล้ายกันจัดไว้ในกลุ่มเดียวกัน ซึ่งการแบ่งกลุ่มที่เกิดจากข้อมูลนี้สามารถตีความเป็นองค์ความรู้ได้

ผลการวิจัยพบว่าผู้ใช้แบบสำรวจสถานการณ์ที่ก่อให้เกิดความเครียดฉบับรูปภาพในรูปแบบคอมพิวเตอร์มีความพึงพอใจกับแบบทดสอบที่ถูกพัฒนาขึ้น และผลการวิจัยยังแสดงให้เห็นอีกว่ามีความเป็นไปได้สูง ที่จะพบองค์ความรู้ที่ซ่อนอยู่ในฐานข้อมูล รวมทั้งแบบสำรวจสถานณการณ์ที่ก่อให้เกิดความเครียดระบบออนไลน์ที่ทำการพัฒนาขึ้น ยังทำให้ผู้ทำการประเมินทราบผลได้ด้วยตนเอง และผู้มีประสบการณ์ด้านจิตวิทยาได้นำข้อมูลไปศึกษาและใช้ประโยชน์ทั้งเชิงวิชาการและการวิจัยต่อไปในอนาคต

88 หน้า

# CONTENTS

# CONTENTS (cont.)

# LIST OF TABLES

# LIST OF FIGURES

# CHAPTER I
# INTRODUCTION

## 1.1 Background and Significance of the Study

For the whole year of 2008, Thailand continuously encountered various problems: political turmoil, economical problems resulted from the late-year hamburger crisis, and so on. As an effect of the economical crisis, the unemployment rate of Thailand in the first quarter of 2009 has been sharply raised for over 100,000 persons. Consequently, many Thai people become stress with unemployment and Financial lean [1]. Stress problem of Thai people seems to be increasingly serious day by day. According to the Ministry of Public Health, statistics from an inventory showed that over 13 percent of Thai people had high or highest levels of stress. Among these stress people, approximately 16 percent of them had notion of hurting themselves, and 9 percent used to hurt themselves. The first four most important causes were found to be financial problems, family problems, sickness or health problems and social/environmental problems [2]. Data from the World Health Organization (WHO) revealed that people throughout the world committed suicide for about 1 million persons per year, which equals to 2,739 persons daily, or 114 persons hourly, or 2 persons for every minute. The report also predicts that, in 2020 number of people who choose to "commit suicide" will rise to 1.5 million persons. In case of Thailand, the Ministry of Public Health reported that the country ranked number 71 in the world for suicide rate [1].

According to statistics, stress is an important psychological problem. People are currently involved with stress activities as their lives are filled with hastiness and competition. In order to fulfill their desire in personal, social, and environmental matter, people are developing pressure to themselves, which eventually leads to stress. Stress is a normal emotional status of our minds that can happen and disappear. However, stress that comes from serious problems might be too serious to handle by oneself.

Therefore, people who notice and understand their own stress will be able to manage situations or adjust themselves properly, and their psychological problems will be reduced. The stress problem can be solved more efficiently if the cause is known and correctly treated. However, since stress is internal condition within a person, many researchers are interested in solutions to assess stress. In Thailand, a form to assess stress problem was developed especially for Thai people by Sucheera Phattharayuttawat et al., which is known as the Life Distress Inventory (LDI) [3]. This tool has been proved to have construct validity and high reliability (Cronbach' Alpha coefficient = 0.87). LDI was developed with a purpose to be usable for general interested people, so that personnel in psychological health field, general practitioners, including people working in health care field can apply it to find information about causes of stress. With ease of use and requiring simple answers, it is also suitable for inventories that require quick results. Assessing stress of people by using LDI will give understanding about the causal problem and level of the stress that will help relieving the stress and lead to optimum solutions.

At present, however, the number of psychological health personnel is relatively low, comparing to the increasing number of psychological patients. According to statistics from the Department of Mental Health for the year 2001, the number of psychiatrists in Thailand was 387, thus ratio of psychiatrists per the whole population equaled 1:161,005 [4]. Hence, psychological patients may have difficulties in accessing to psychiatrists, and the treatment service might be delayed. Computer technology, therefore, should be applied to enhance and facilitate medical services for psychological patients. Computer has played a major role in enhancing efficiency and efficacy of many works, especially regarding medicine and public health. With its capability in various applications, computer has become an important tool that can facilitate many medical activities. Some hospitals used computer for collecting database of patients, diagnose symptoms, and cure diseases. Currently, psychological assessment has begun using computer for many applications including test procedures, scoring, result interpretation, as well as analyzing items in questionnaires statistically.

In addition, computer can store all data in the database for future analyses or development of tests [5].

The researcher is therefore interested in developing stress tests in the form of web application, which is a technology for developing applications on the internet, with capabilities to update and manage data. Web applications are not required to be installed in user's computer. Users can use the application via the internet by using their preferred web browser. The internet is used for this study since it currently plays a crucial role in the world of information technology by providing a great amount of information to people who are globally allowed to search for information and utilize many online services without a need for transportation. For this research, the Life Distress Inventory (LDI) was utilized to be presented in the form of web application, with an objective to increase public awareness about stress problem. Web application technology is considered highly appropriate for the research's objectives since it has been continuously developed and widely used. In addition, since more and more people can access to the internet, the developed web application for stress can serve a great number of patients.

## 1.2 Research objectives

1. To develop an online Web questionnaires database system for measure of life distress inventory. The report can be generating for clinician and researcher.

2. To discover some knowledge hidden in the life distress inventory database.

## 1.3 Research tools

1. The web-application form of the Life Distress Inventory (LDI)
2. The standard form of Life Distress Inventory (LDI)
3. Questionnaires for assessing satisfaction of people toward utilization of the developed web-application form of the Life Distress Inventory (LDI)

## 1.4 Scope of work

1. The aim of this research is to design and implement web questionnaire database for The Life Distress Inventory questionnaires.

2. To study and design database for collecting the scores, and measuring level of rater.

3. To consider database management with database relational.

4. To consider documents with test-based document such as ASP.NET, HTML.

5. To design Database by using Microsoft SQL.

6. Use clustering and decision tree for knowledge discovery.

## 1.5 Expected outcome

1. Utilization of psychological questionnaire via the format of web application will enhance attractiveness to users, and increase their convenience by allowing them to answer the questionnaire wherever they can access to the internet.

2. By using a web application of questionnaire to collect data regarding psychological health, psychological personnel can have convenience in using the collected data for analyzing causes of stress, since the web application will always store data of users in its database system.

3. It is expected that the developed web application will be further used in developing other tools.

# CHAPTER II
# LITERATURE REVIEW

The objectives this research is to developed Online Rating Scale For The Life Distress Inventory and discover some knowledge hidden in the life distress inventory database. The reviewed topics and researches are as follows.

## 2.1 Stress

### 2.1.1 Definitions of stress

Originally, the word stress was term used in physics, primarily to describe enough tension or force placed on an object to bend or break it.The word stress as applied to humen condition was first made popular by noted physiologist Hans Selye in his book The Stress of Life, where he described his research: to understand the physiological responses to chronic stress and it relationship to disease. Today, the word stress is frequently used to describe the level of tension people feel is placed on their minds and souls by the demands of their jobs, relationships, and responsibilities in their personal lives [6].

Hans Selye's [6] define stress is the nonspecific response of the body to any demand placed upon it to adapt, whether that demand produces pleasure or pain.

A Holistic Medicine [6] define stress is the inability to cope with a perceived or real ( or imagined) threat to one's mental, physical, emotional ,and spiritual well–being ,which results in a series of physiological responses and adaptation.

Lazarus & Folkman [7] have defined stress as follows:

1. Response definitions of stress

Stress means individuals' reactions to stimuli which are seen as dangerous or threatening. This can be measured from physiological responses to stress such as faster heartbeat, trembling hands, indigestion, restlessness, muscle spasm, and sweating.

2. Stimulus definitions of stress

Stress means individuals' encounter with major changes in life which bring about stress such as death of loved ones, being laid off from work, or divorce.

3. Person-environment relationship definitions of stress

Stress means individuals' interaction with their environment and their cognitive appraisal of such environment. When individuals appraise that their resources are not sufficient and need to pull off reserved resources, they can develop stress. In fact,this is what Lazarus uses to define stress.

**2.1.2 Type of stress**

There are three kinds of stress :eustress,neustress and distress [6].

1. Eustress is good stress and arises in any sitution or cumstance that person finds motivating or in spiring such as falling in love might be an exanple of eustress , meeting a movie star or professional athlete may also be a type of eustress.Usually, situtions that are classified as eustress are enjoyable and for this reason are not considered to be a threat.

2. Neustress describes sensory stimuli that have no consequential effect; it is considered neither good or bad such as news of an earthquake in a remote corner of the world.

3. Distress is considered bad and often is abbreviated simply as stress. There are two kinds of distress

- Acute stress : Stress that is intense in nature but short in duration

- Chronic stress: Stress that is not as intense as acute stress but that linger for prolonged periods of time (e.g.,hours, days, weeks, or months ) such as financial problem.

**2.1.3 Causes of stress**.

Wallace [8] divided causes of stress into 2 types.

1. External stressor was the person's extrinsic cause mostly involving the interpersonal relationship problem such as family relationship between spouse or parents and children, relationship in the work place between employers and employees, colleagues, work progression, encouragement or connection with other people such as a conflict of opinions and interests, etc.

2. Internal stressor was the person's intrinsic cause involving pain, sickness, cognition, guilt, dream, expectation, ambition and self concept.

Panpreecha, C. [9] classified the stressors as follows:

1. Internal cause was the stressors that came from inside the person including 2 categories. The first was physical factor such as physical conditions that produced stress. Because the body and mind cannot be separated, when the body was stressful, so did the mind. These conditions included a lack of relaxation, physical illness, and alcohol or drug dependence. The other cause was psychological factor involving some mental conditions that produced stress such as negative emotion, anger, fear, frustration, impulsiveness and dependent need etc.

2. External cause was the stressor that came from outside the person including as follows:

2.1 Loss of beloved person, property, work, being dismissed, reduced or transferred position, and bankruptcy.

2.2 Life change or the turning point of life resulted in psychological imbalance. It included first time working, marriage, having first child, experiencing the menopause, changing environment that needed adaptation, and sudden change without expectation and preparation such as work or home moving.

2.3 Several dangers that threatened lives and properties included the living in the place with a lot of thieves, being threatened, flooding and fire. These states produced severe mental pressure that can result in extreme stress.

2.4 Stressful working included a lack of security in working, risky job, unsatisfactory job and considering it unimportant.

2.5 Low economic status included being poor or in debt resulting in incomplete growth, nutrition deficiency, and living in crowded community without security resulting in anxiety, fear, pressure and stress.

2.6 Social status included the people's moving into the urban, the crowd, competition, inconvenient traveling because of the traffic, not warm family, and lack of fresh air.

### 2.1.4 Level of Stress

Janis [10] divides stress in 3 levels as follows:

1. Low Stress Level. The stress will end shortly for a second or an hour only.The level involves with tiny reasons such as traffic jam, missing the transportation,late arrival.

2. Medium Stress Level. The level is more severe than the low level. It may last for hours or days. It may cause some sickness after the situation past such as the opposition with colleague.

3. High Stress Level. The stress will stay for weeks or months or years such as death of dear people, serious illness, lost of important organs for living normal life.

Department of Mental Health [11] which classifies stress level in 5 as follows:

1. Very Below Average Stress Level. Individual is able to manage daily life stress in stage which has less pressure or temptation in living life than others.

2. Normal Stress Level. Individual is able to manage stress in daily life and adjust to all situations suitably in this stage.

3. Mildly Above Average Stress Level. At this stage, individual feels uncomfortably from daily life problems or unsolved obstructions or arguments which is the stress facing in daily life.

4. Medium Above Average Stress Level. Individual starts to feel tight and high stress at this stage and very worry about emotional problems from arguments and crisis in life. It can be noticed

5. Extremely Above Average Stress Level. Individual starts to feel serious stress or in crisis. If stress is continue without proper and suitable solution, it may lead to serious illness which harm individual and close people.

### 2.1.5 Measurement of Stress

Adolf Meyer [12] suggested that psychiatrists use a life chart in evaluating their patients. The idea is elegant in its simplicity. One plots a patient's life on a line, beginning with birth and ending in the present. One then pencils in when various important life events have occurred and also indicates the onset or worsening of health problems. The life chart turns out to be a very effective way of eliciting a clinical history. Sometimes, recurrent patterns of association between life events and illness episodes are suggested.

Holmes and Rahe [13] studied and developed life-event scale, especially the Social Readjustment Rating Scale (SRRS) to analyze the patient stress severity levels. It collects all information about patient’s illness or onset and classifies the stress severity levels in order that doctor will know about all stress types having effect on patient’s mind from the minimum to maximum. 10 severe stress levels are as follows:

1. Death of spouse
2. Divorce
3. Marital separation
4. Jail term
5. Death of close family member
6. Personal injury or illness
7. Marriage
8. Fired at work
9. Marital recociliation
10. Retirement

10 medium severe stress levels are as follows:

1. Pregnancy
2. Sex difficulties
3. Gain of new family member
4. Business readjustment
5. Change in financial state
6. Death of close friend
7. Change to different line of work
8. Change in number of argument with spouse
9. Large amount of Mortgage
10. Change in responsibility at work

## 2.2 The Life Distress Inventory

The Life Distress Inventory (LDI) was created by Sucheera Phattharayuttawat, Tienchai Ngamthipwattana, Kanokrat Sukhatungka, Somporn Charascharoenwittaya, and Malai Chaloemchainukul of Department of Psychiatry, Faculty of Medicine Siriraj Hospital, Mahidol University [3]. Their objective was to develop an inventory form that can be used to assess distress that happens in daily life of Thai people. This inventory form features these following aspects of efficiency.

1. Construct validity
2. Reliability
3. Sensitivity and specificity
4. Having norms as criteria for selecting and eliminating causal events.

The Life Distress Inventory form has efficiency in assessing distress and its causes. It is easy to use and to understand. The testing method is not complex or complicated. The LDI form allows personnel in psychological health field, general practitioners, or people working in health care field to find information about causes of stress for correct or proper treatment.

Development of the Life Distress Inventory

Step 1 : Create questionnaire by developing questions derived from reviewing literature regarding situations that cause distress in daily life of general people.

Step 2 : Use the questionnaire created in Step 1 to test with a sampling group of 200 individuals that are classified as either normal people or psychological patients.

There were 20 items altogether that were created. For factor analysis, the items were analyzed by using Varimax rotation. Only items with factor loadings exceeding 0.40 were selected. Factor loadings of all items are summarized in Table 2.1.

**Table 2.1** The Varimax rotation of first factors by according for using a 0.40 as a criterion

| **Questions comprising items with less than 0.40** | **Questions comprising item with more than 0.40** |
|---|---|
| 20 | 1 , 2 , 3 , 4 , 5 , 6 , 7 , 8 , 9 , 10 , 11 , 12<br>13 , 14 , 15 , 16 , 17 , 18 , 19 |

From analyzing the items to find Construct validity of the questionnaire, it was found that the items can be classified into 5 factors, of which the variance accounted for 50.12% of the total variance. The 5 factors were found to have eigenvalues more than 1, as shown in Table 2.2.

**Table 2.2** Unrotated principal axes analyzed of the life distress inventory

| **Factor No.** | **Eigenvalue** | **Name of Factor** |
|---|---|---|
| 1 | 21.8416 | Marriage |
| 2 | 18.0531 | Education and occupation |
| 3 | 11.5287 | Activity outside family |
| 4 | 8.2483 | Personal life and family |
| 5 | 8.0014 | Satisfaction in life |

After rotating axes of the 5 factors, they were found to have variance for 50.12% of the total variance. Items that belong to each factor are listed in Table 2.3.

**Table 2.3** Varimax rotation of five factor accounting for 50.12% of variance

| Factors | Listed questions | Item No. |
|---|---|---|
| 1 | Marriage | 1 , 2 , 3 |
| 2 | Education and occupation | 8 , 9 , 10,12 |
| 3 | Activity outside family | 11 , 18 , 19 |
| 4 | Personal life and family | 4 , 5 , 6 , 7 ,13 ,14 |
| 5 | Satisfaction in life | 15 , 16 ,17 |

Factor loadings (weight) of all items are shown in Table 2.4.

**Table 2.4** Weight of each item in factor

| Item | Events | Weight |
|---|---|---|
| 1 | Marriage | .4102 |
| 2 | Gender adaptation | .4520 |
| 3 | Relationship with spouse /love affair | .5012 |
| 4 | Relationship with kids | .5086 |
| 5 | Relationship with relatives | .4310 |
| 6 | General living in home | .4711 |
| 7 | Financial status | .5012 |
| 8 | Education | .6125 |
| 9 | Unemployed | .4417 |
| 10 | Atmosphere and surrounding at office | .5812 |
| 11 | Leisure time | .4698 |
| 12 | Time management | .5542 |
| 13 | Physical health | .4369 |
| 14 | Mental health | .4818 |
| 15 | Freedom in personal life | .5312 |
| 16 | Satisfaction in life | .5587 |
| 17 | Future anticipation | .4439 |
| 18 | Adaptation to friends | .6019 |
| 19 | Social adaptation | .5103 |

The analysis for reliability was conducted by using Test–retest methodology of Pearson correlation analysis. Additionally, analysis for internal consistency of reliability was conducted by calculating for Cronbach alpha coefficients. Results are shown in Table 2.5 and Table 2.6.

**Table 2.5** Test-retest reliability Coefficients

| Listed questions | Test-retest Coefficients |
|---|---|
| 1. Marriage | .77 |
| 2. Education and occupation | .88 |
| 3. Activity outside family | .87 |
| 4. Personal life and family | .82 |
| 5. Satisfaction in life | .84 |
| Total | .83 |

**Table 2.6** Reliability coefficients by Cronbach's Alpha

| Listed questions | Cronbach's Alpha Coefficients |
|---|---|
| 1. Marriage | .89 |
| 2. Education and occupation | .85 |
| 3. Activity outside family | .91 |
| 4. Personal life and family | .85 |
| 5. Satisfaction in life | .82 |
| Total | .87 |

**Table 2.7** Mean and standard deviation (SD) classified by questioning items

| Listed questions | Mean | SD | P-value |
|---|---|---|---|
| 1. Marriage | 2.87 | 0.89 | .000 |
| 2. Education and occupation | 2.19 | 1.11 | .000 |
| 3. Activity outside family | 2.55 | 0.77 | .000 |
| 4. Personal life and family | 2.82 | 1.59 | .000 |
| 5. Satisfaction in life | 2.11 | 0.66 | .000 |
| Average total | 2.17 | 0.72 | .000 |

From Table 2.7, the results were from the analysis for checking Discriminant validity of the questionnaire and each item in the questionnaire by examining scores of all questionnaires given to the group of normal people and the group of distress patients. It was found that scores of each item and total score from the questionnaire between the normal group and the distress group were statistically different, at a significant level of .001.

In order to use the questionnaire for selecting or eliminating situations that cause distress in daily life of people, the analyses for sensitivity and specificity were conducted. Consideration will be based on the optimum crossing point, which was found to be 18 with a sensitivity of .85 and a specificity of .74.

**Purposes to use the questionnaire (the inventory form)**

The questionnaire can be used to examine situations that cause distress in regular daily life. The questionnaire originally comprised 19 usable items. One more item was added, thus it now has 20 questioning items in total.

**Scoring**

Total score of the questionnaire ranged from 0 to 80. The higher score represent high distress. Scoring of each item can be classified can be classified by factors as the following.

**Table 2.8** Listed question substituting of each factor

| Factors | Item No. |
|---|---|
| Marriage | 1 , 2 , 3 |
| Education and occupation | 8 , 9 , 10 , 12 |
| Activity outside family | 11, 18 , 19 |

**Table 2.8** Listed question substituting of each factor (continued)

| Factors | Item No. |
|---|---|
| Personal life and family | 4 , 5 ,6 , 7 , 13, 14 |
| Satisfaction in life | 15 , 16 , 17 ,20 |

**Norms**

Scores of situations or events in daily life that cause stress can be categorized into 5 levels, as shown in Table 2.9.

**Table 2.9** Score of events causing stress in daily life in each level.

| Score of events causing stress in daily life | Stress level |
|---|---|
| 34-80 | Cause severe stress |
| 26-33 | Cause moderate stress |
| 19-25 | Cause little stress |
| 4-18 | Normal stress |
| 0-3 | Very below average stress |

## 2.3 SQL [14]

Every business has data, which requires some organized method or mechanism for maintaining the data. This mechanism is referred to as a *database management system* (*DBMS*). Database management systems have been around for years, many of which started out as flat-file systems on a mainframe. With today's technologies, the accepted use of database management systems has begun to flow in other directions, driven by the demands of growing businesses, increased volumes of corporate data, and of course, Internet technologies.

The modern wave of information management is primarily carried out through the use of a *relational database management system* (*RDBMS*), derived from the traditional *DBMS*. Modern databases combined with client/server and Web technologies are typical combinations used by current businesses to successfully manage their data and stay competitive in their appropriate markets. The trend for many businesses is to move from a client/server environment to the Web, where location is not a restriction when users need access to important data. The next few sections discuss SQL and the relational database, the most common *DBMS* implemented today. A good fundamental understanding of the relational database, and how to apply SQL to managing data in today's information technology world, is important to your understanding of the SQL language.

### 2.3.1 Definition of SQL

SQL, Structured Query Language, is the standard language used to communicate with a relational database. The prototype was originally developed by IBM using Dr. E.F. Codd's paper ("A Relational Model of Data for Large Shared Data Banks") as a model. In 1979, not long after IBM's prototype, the first SQL product, ORACLE, was released by Relational Software, Incorporated (it was later renamed Oracle Corporation). It is, today, one of the distinguished leaders in relational database technologies. SQL is pronounced either of two ways: as the letters S-Q-L, or as "sequel"; both pronunciations are acceptable. However, most experienced SQL users tend to use the latter pronunciation.

The American National Standards Institute (ANSI) is an organization that approves certain standards in many different industries. SQL has been deemed the standard language in relational database communication, originally approved in 1986 based on IBM's implementation. In 1987, the ANSI SQL standard was accepted as the international standard by the International Standards Organization (*ISO*). The standard was revised again in 1992 and was called SQL-92. The newest standard is now called SQL-99; it's also referred to as SQL3.

**The New Standard: SQL-99**

SQL-99 has five interrelated documents and other documents may be added in the near future. The five interrelated parts are as follows:

- *Part 1—SQL/Framework*— Specifies the general requirements for conformance and defines the fundamental concepts of SQL.
- *Part 2—SQL/Foundation*— Defines the syntax and operations of SQL.

• *Part 3—SQL/Call-Level Interface*— Defines the interface for application programming to SQL.

• *Part 4—SQL/Persistent Stored Modules*— Defines the control structures that then define SQL routines. Part 4 also defines the modules that contain SQL routines.

• *Part 5—SQL/Host Language Bindings*— Defines how to embed SQL statements in application programs that are written in a standard programming language.

### 2.3.2 Definition of Database

In very simple terms, a *database* is a collection of data. Some like to think of a database as an organized mechanism that has the capability of storing information, through which a user can retrieve stored information in an effective and efficient manner.

People use databases every day without realizing it. A phone book is a database. The data contained consists of individuals' names, addresses, and telephone numbers. The listings are alphabetized or indexed, which allows the user to reference a particular local resident with ease. Ultimately, this data is stored in a database somewhere on a computer. After all, each page of a phone book is not manually typed each year a new edition is released.

The database has to be maintained. As people move to different cities or states, entries may have to be added or removed from the phone book. Likewise, entries will have to be modified for people changing names, addresses, or telephone numbers, and so on. Figure 2.1 illustrates a simple database.

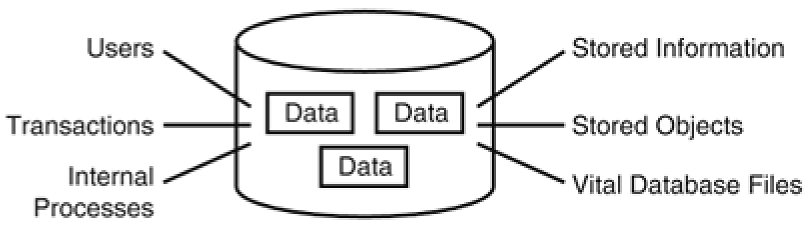



**Figure 2.1** The database.

### 2.3.3 Relational Database

A *relational database* is a database divided into logical units called *tables,* where tables are related to one another within the database. A relational database allows data to be broken down into logical, smaller, manageable units, allowing for easier maintenance and providing more optimal database performance according to the level of organization. In Figure 2.2, can see that tables are related to one another through a common key (data value) in a relational database.

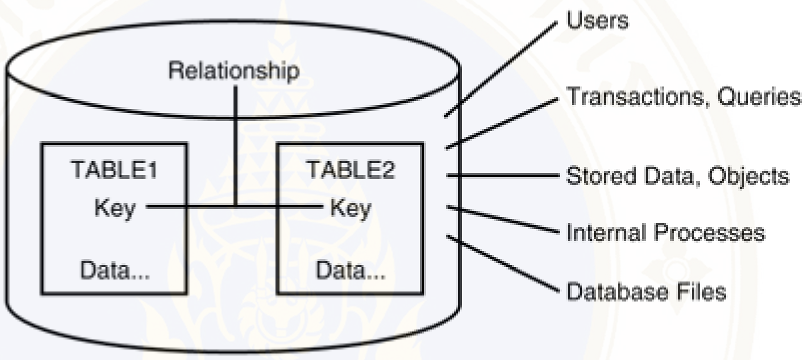



**Figure 2.2** The relational database.

Again, tables are related in a relational database, allowing adequate data to be retrieved in a single query (although the desired data may exist in more than one table). By having common *keys,* or *fields,* among relational database tables, data from multiple tables can be joined to form one large result set.

*A relational database* is a database composed of related objects, primarily tables. A *table* is the most basic means of storage for data in a database.

### 2.3.4 Client/Server Technology

In the past, the computer industry was predominately ruled by mainframe computers; large, powerful systems capable of high storage capacity and high data processing capabilities. Users communicated with the mainframe through dumb terminals—terminals that did not think on their own, but relied solely on the mainframe's CPU, storage, and memory. Each terminal had a data line attached to the mainframe. The mainframe environment definitely served its purpose, and does today in many businesses, but a greater technology was soon to be introduced: the client/server model.

In the *client/server system,* the main computer, called the *server,* is accessible from a network—typically a local area network (*LAN*) or a wide area network (*WAN*). The server is normally accessed by personal computers (PCs) or by other servers, instead of dumb terminals. Each PC, called a *client*, is provided access to the network, allowing communication between the client and the server, thus explaining the name client/server. The main difference between client/server and mainframe environments is that the user's PC in a client/server environment is capable of thinking on its own, capable of running its own processes using its own CPU and memory, but readily accessible to a server computer through a network. In most cases, a client/server system is much more flexible for today's overall business needs and is much preferred.

Modern database systems reside on various types of computer systems with various operating systems. The most common types of operating systems are Windows-based systems and common line systems such as UNIX. Databases reside mainly in client/server and Web environments. A lack of training and experience is the main reason for failed implementations of database systems. Nevertheless, an understanding of the client/server model and Web-based systems is imperative with the rising (and sometimes unreasonable) demands placed on today's businesses as well as the development of Internet technologies and network computing. Figure 2.3 illustrates the concept of client/server technology.

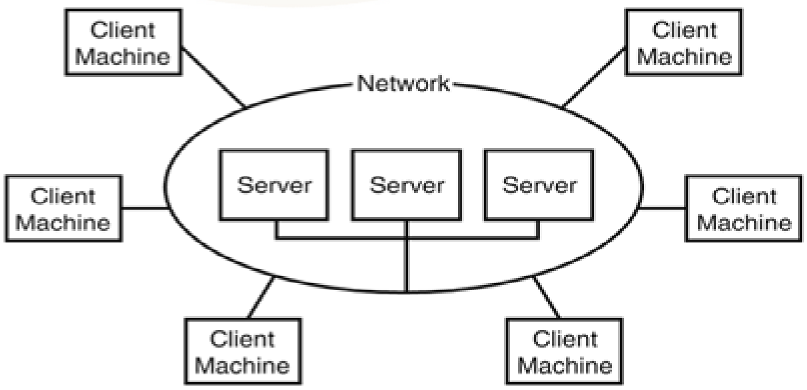



**Figure 2.3** The client/server model.

### 2.3.5 Web-Based Database Systems

Business information systems are moving toward Web integration. Databases are now accessible through the Internet, meaning that customers' access to an organization's information is enabled through an internet browser such as Internet Explorer or Netscape. Customers (users of data) are able to order merchandise, check on inventories, check on the status of orders, make administrative changes to accounts, transfer money from one account to another, and so forth.

A customer simply invokes an Internet browser, goes to the organization's Web site, logs in (if required by the organization), and uses an application built into the organization's Web page to access data. Most organizations require users to register with them, and will issue a login and password to the customer.

Of course, there are many things that occur behind the scenes when a database is being accessed via a Web browser. SQL, for instance, can be executed by the Web application. This executed SQL is used to access the organization's database, return data to the Web server, and then return that data to the customer's Internet browser.

The basic structure of a Web-based database system is similar to that of a client server system from a user's standpoint. Refer to Figure 2.3. Each user has a client machine, which has a connection to the Internet and contains a Web browser. The network in Figure 2.3 (in the case of a Web-based database) just happens to be the Internet, as opposed to a local network. For the most part, a client is still accessing a server for information. It doesn't matter that the server may exist in another state, or even another country. The main point of Web-based database systems is to expand the potential customer base of a database system that knows no physical location bounds, thus increasing data availability and an organization's customer base.

## 2.4 ASP.NET

ASP.NET [15] is a web application framework developed and marketed by Microsoft to allow programmers to build dynamic web sites, web applications and web services. It was first released in January 2002 with version 1.0 of the .NET

Framework, and is the successor to Microsoft's Active Server Pages (ASP) technology. ASP.NET is built on the Common Language Runtime (CLR), allowing programmers to write ASP.NET code using any supported .NET language. The ASP.NET SOAP (Simple Object Access Protocol)extension framework allows ASP.NET components to process SOAP message.

ASP.NET is a technology for developing web applications. A web application is just a fancy name for a dynamic web site. Web applications usually (but not always) store information in a database, and allow visitors to the site to access and change that information. Many different programming technologies and supported languages have been developed to create web applications; PHP, JSP, Ruby on Rails, CGI, and ColdFusion are just a few of the more popular ones.

ASP.NET uses the Microsoft .NET Framework. The .NET Framework collects all the technologies needed for building Windows desktop applications, web applications, web services, and so on, into a single package, and makes them available to more than 40 programming languages.

ASP.NET has a few unique features:

• ASP.NET lets write the server-side code using your favorite programming language—or at least one the one you prefer from the long list of supported languages. The .NET Framework currently supports over 40 languages, and many of these may be used to build ASP.NET web sites. The most popular choices are C# (pronounced “C sharp”) and Visual Basic (or VB).

• ASP.NET pages are compiled, not interpreted. In ASP.NET’s predecessor, ASP, pages were interpreted: every time a user requested a page, the server would read the page’s code into memory, figure out how to execute the code, and execute it. In ASP.NET, the server need only figure out how to execute the code once. The code is compiled into efficient binary files, which can be run very quickly, again and again, without the overhead involved in rereading the page each time. This allows a big jump in performance, compared to the old days of ASP.

• ASP.NET has full access to the functionality of the .NET Framework. Support for XML, web services, database interaction, email, regular expressions, and many other technologies are built right into .NET, which saves you from having to reinvent the wheel.

- ASP.NET allows you to separate the server-side code in your pages from the HTML layout. When you're working with a team composed of programmers and design specialists, this separation is a great help, as it lets programmers modify the server-side code without stepping on the designers' carefully crafted HTML—and vice versa.
- ASP.NET makes it easy to reuse common User Interface elements in many web forms, as it allows us to save those components as independent web user controls.
- You can get excellent tools that assist in developing ASP.NET web applications. Visual Web Developer 2008 is a free, powerful visual editor that includes features such as code autocompletion, code formatting, database integration functionality, a visual HTML editor, debugging, and more. In the course of this book, you'll learn how to use this tool to build the examples we discuss.
- The .NET Framework was first available only to the Microsoft Windows platform, but thanks to projects such as Mono, it's since been ported to other operating system.

## 2.5 Data mining

Data mining, also known as knowledge discovery in databases (KDD), is a process that uses statistical, mathematical, artificial intelligence and machine-learning techniques to identify and extract useful knowledge from a large amount of data. While data mining and knowledge discovery in database (or KDD) are frequently treated as synonyms, data mining is actually a part of the knowledge discovery process[16].

The KDD process comprises a number of steps ranging from the collection of raw data to the extraction of new knowledge. The iterative process consists of the following steps :

- **Data cleaning :** also known as data cleansing. It is a phase in which noises and irrelevant data are removed from the collection.
- **Data integration :** In this step, multiple data sources, often heterogeneous, may be combined into a single common source.
- **Data selection :** In this step, the data relevant to the analysis is identified and retrieved from the data collection.
- **Data transformation :** also known as data consolidation. It is a phase in which the selected data is transformed into the form appropriate for the mining procedure.
- **Data mining :** It is the crucial step in which a number of techniques are applied to extract potentially useful patterns.
- **Pattern evaluation :** In this step, patterns representing interesting knowledge are identified based on given measures.
- **Knowledge representation :** It is the final phase in which the discovered knowledge is visually presented to users. This step uses visualization techniques to help the users understand and interpret the results.

**Data Mining Tasks and Techniques** [17]

Data mining tasks are often divided into two major categories:

**Predictive** The goal of predictive tasks is to use the values of some variables to predict the values of other variables. For example, in Web mining, e-tailers are interested in predicting which online users will make a purchase at their Web site. Other examples include biologists, who would like to predict the functions of proteins, and stock market analysts, who would like to forecast the future prices of various stocks.

**Descriptive** The goal of descriptive tasks is to find human-interpretable patterns that describe the underlying relationships in the data. For example, Earth Scientists are interested in discovering the primary forcings influencing observed climate patterns. In network intrusion detection, analysts want to know the kinds of cyber-attacks being launched against their networks. In document analysis, it is useful to find groups of documents, where the documents in each group share a common topic.

Data mining tasks can be accomplished using a variety of data mining techniques, as shown in Figure 2.4.

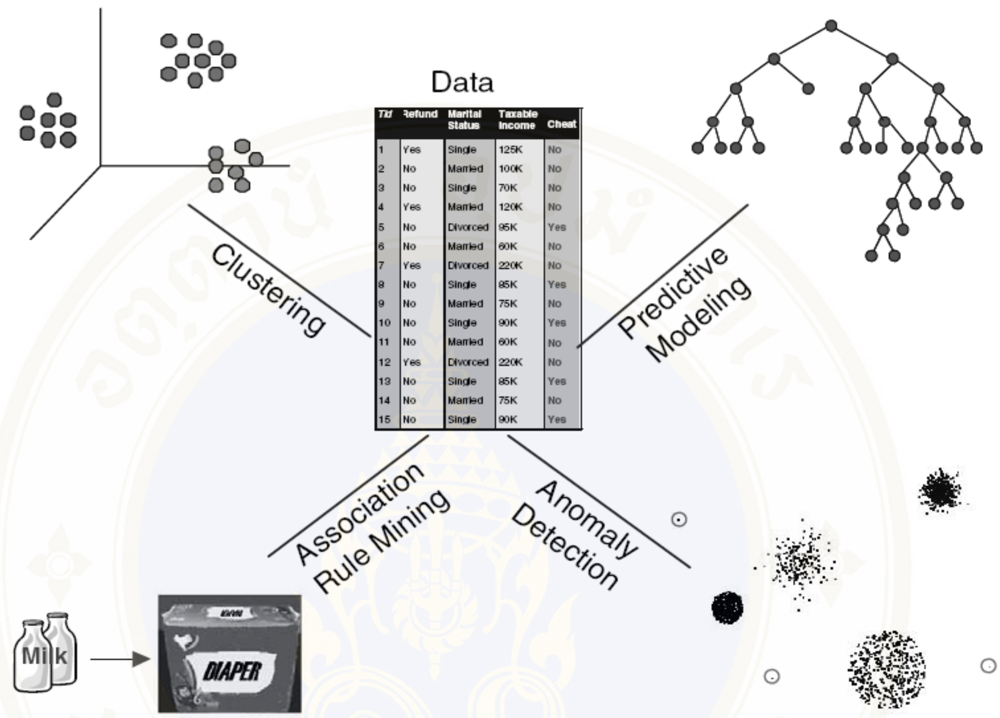



**Figure 2.4** : Data mining techniques.[17]

• **Predictive modeling** is used primarily for predictive data mining tasks. The input data for predictive modeling consists of two distinct types of variables: (1) explanatory variables, which define the essential properties of the data, and (2) one or more target variables, whose values are to be predicted. For the Web mining example given in the previous section, the input variables correspond to the demographic features of online users, such as age, gender, and salary, along with their browsing activities, e.g., what pages are accessed and for how long. There is one binary target variable, Buy, which has values, Yes or No, indicating, respectively, whether the user will buy anything from the Web site or not. Predictive modeling techniques can be further divided into two categories: *classification* and *regression*. Classification techniques are used to predict the values of discrete target variables, such as the Buy variable for online users at a Web site. For example, they can be used to predict whether a customer will most likely be lost to a competitor, i.e., customer churn or attrition, and to determine the category of a star or galaxy for sky survey cataloging. Regression techniques are used to predict the values of continuous target variables, e.g., they can be applied to forecast the future price of a stock.

• **Association rule mining** seeks to produce a set of dependence rules that predict the occurrence of a variable given the occurrences of other variables. For example, association analysis can be used to identify products that are often purchased together by sufficiently many customers, a task that is also known as *market basket analysis*. Furthermore, given a database that records a sequence of events, e.g., a sequence of successive purchases by customers, an important task is that of finding dependence rules that capture the temporal connections of events. This task is known as *sequential pattern analysis*.

• **Cluster analysis** finds groupings of data points so that data points that belong to one cluster are more similar to each other than to data points belonging to a different cluster, e.g., clustering can be used to perform market segmentation of customers, document categorization, or land segmentation according to vegetation cover. While cluster analysis is often used to better understand or describe the data, it is also useful for summarizing a large data set. In this case, the objects belonging to a single cluster are replaced by a single *representative* object, and further data analysis is then performed using this reduced set of representative objects.

• **Anomaly detection** identifies data points that are significantly different than the rest of the points in the data set. Thus, anomaly detection techniques have been used to detect network intrusions and to predict fraudulent credit card transactions. Some approaches to anomaly detection are statistically based, while other are based on distance or graph-theoretic notions.

## 2.6 Clustering

A cluster is a group of objects that are similar to one another within the cluster, and dissimilar to objects in the other clusters. Clusters should be homogenous and separable, i.e. all instances within a cluster should be relatively similar (homogenous). Each cluster should be very different from the other clusters (heterogeneous) [18].The general objective of clustering is to arrive at a certain number of (initially unnamed) clusters and to assign descriptions which characterize each of them.

Clustering focuses on descriptive tasks on historic data. Nevertheless, good clusters can also be useful for prediction, as one can :

- Create clusters
- Assign a new instance to the closest cluster and predict its properties using the properties of that cluster.

In the diagram below (Figure 2.5), there are four clusters of car purchasers, each characterized by three attributes : income, number of children, and type of car [19].

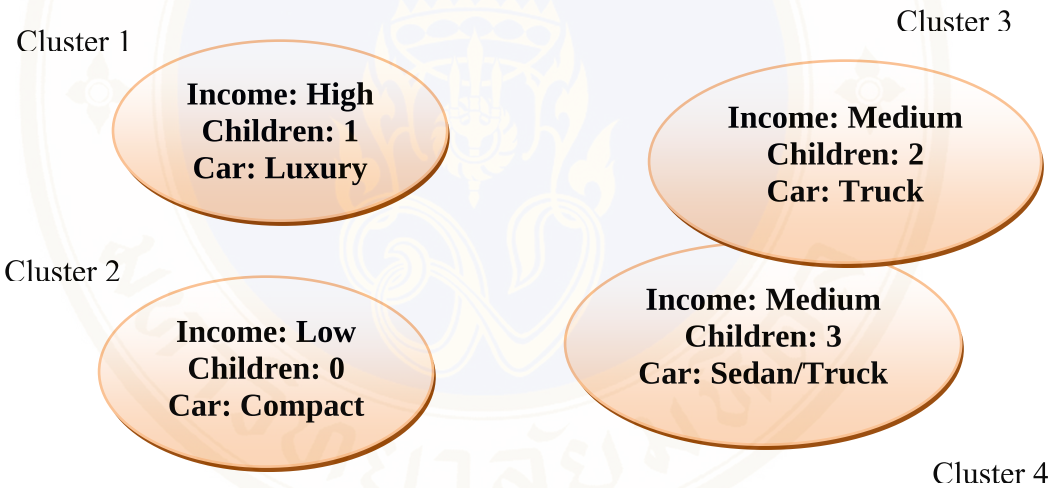



**Figure 2.5** Four clusters of car purchasers [19]

Notice that these clusters provide a concise description of the dataset. A typical follow-up analysis is to investigate the similarities and the differences between each cluster.

### 2.6.1 Cluster analysis

Cluster analysis is an exploratory data analysis method for solving classification problems. The importance of cluster analysis is that it may reveal associations and structures in data that were not previously apparent but are sensible and useful once found.

**Results from the analysis may be used to**

1) Suggest statistical models to describe the populations

2) Help identify a classification scheme (e.g., types of customers)

3) Find typical cases to represent each class, according to the classification scheme

4) Derive rules for assigning any new case to one of these classes for identification, targeting, and diagnostic purposes

5) Provide measures of definition, size, and change in what were previously broad concepts

**Cluster analysis methods**

Cluster analysis may be based on one or more of the following, general methods :

1) Statistics (including both hierarchical and nonhierarchical)

2) Optimization

3) Neural networks

4) Fuzzy logic

5) Genetic algorithms

Each of these methods follow either one of the two approaches :

- Divisive : all items start in one cluster and are broken apart.
- Agglomerative : all items start in individual clusters, and the clusters are joined together.

Most cluster analysis methods involve the use of a distance between pairs of items. That is, there is a measure of similarity between every pair of items to be clustered. Weighted averages may be used to establish these distances.

### 2.6.2 k-Means Clustering

K-means clustering [20] is a straightforward and effective algorithm for finding clusters in data. The algorithm proceeds as follows :

- **Step 1 :** Ask the user how many clusters $k$ the dataset should be partitioned into.

- **Step 2 :** Randomly assign *k* records to be the initial cluster center locations.
- **Step 3 :** For each record, find the nearest cluster center. Thus, in a sense, each cluster center "owns" a subset of the records, thereby representing a partition of the dataset. The *k* clusters are *C1*, *C2*, …, *Ck*.
- **Step 4 :** For each of the *k* clusters, find the cluster *centroid*, and update the location of each cluster center to the new value of the centroid.
- **Step 5 :** Repeat steps 3 to 5 until convergence or termination.

The "nearest" criterion in step 3 is usually Euclidean distance, although other criteria may be applied as well. The cluster centroid in step 4 is found as follows.Suppose that there are *n* data points (*a1, b1, c1*), (*a2, b2, c2*), …, (*an, bn, cn*), the *centroid* of these points is the center of gravity which is located at point (*Σai/n, Σbi/n, Σci/n*).

For example, the points (1, 1, 1), (1, 2, 1), (1, 3, 1), and (2, 1, 1) would have the centroid

$$\frac{1+1+1+2}{4}, \frac{1+2+3+1}{4}, \frac{1+1+1+1}{4} = (1.25, 1.75, 1.00)$$

The algorithm terminates when the centroids no longer change. In other words, if terminates when, for all clusters *C1*, *C2*, …, *Ck*, all the records "owned" by each cluster center remain in that cluster. Alternatively, the algorithm may terminate when some convergence criterion is met, such as no significant shrinkage in the *sum of squared errors* :

$$SSE = \sum_{i=1}^{k} \sum_{p \in Ci} d(p, m_i)2$$

where $p \in Ci$ represents each data point in cluster *i* and $m_i$ represents the centroid of cluster *i* [21].

## 2.7 Classificaion

Classification [22] is perhaps the most often employed data mining technique. It involves a set of instances or predefined examples to develop a model that can classify the population of records at large.

The use of classification algorithms begins with a sample set of preclassified example transactions. For a fraud detection application, this would include complete records of both fraudulent and valid transactions, determined on a record-by-record basis. The classifier-training algorithm uses these preclassified examples to determine the set of parameters required for proper identification. The algorithm then encodes these parameters into a model called a classifier, or classification model. The approach affects the decision-making capability of the system. Once an effective classifier is developed, it is used in a predictive mode to classify new records automatically into these same predefined classes.

In the fraud detection case cited above, the classifier would be able to identify probable fraudulent activities. Another example would involve a financial application in which a classifier capable of identifying risky loans could be used to aid in the decision of whether or not to grant a loan to an individual.

### 2.7.1 Classification analysis

Classification analysis is the organization of data in given classes. Also known as supervised classification, it uses the given class labels to order the objects in the data collection. The analysis is normally performed on a training set where all objects are already associated with known class labels. The classification algorithm learns from the training set and builds a model. The model is then used to classify new objects.

### 2.7.2 Construction of C4.5 Decision Trees

C4.5 is a classification technique that gives decision trees as results. A C4.5 decision tree has the following properties: leaf nodes represent target classes or possible values of the target attribute; internal nodes represent independent attributes; and arcs leaving an internal node represent decisions according to attribute values of that node. C4.5 works with nominal data [23]. Its algorithm is roughly as follows: within each iteration, an independent attribute that best classifies instances will be selected to be a new internal node. Suppose that this node is called the current node.A set of child nodes will be generated along with arcs representing decision paths from the current node to its children. The algorithm stops once all the child nodes successfully classify instances to the target classes. Otherwise, another independent attribute will be selected and put into a wavering node (which becomes the new current node). The best classifying attribute is selected on the basis of its yielding information gain ratio. That is, entropy before and after the classification by each attribute will be measured. The most reduction of entropy indicates the least energy remaining in the data, or the most stability incurred by the classification, hence the most information gain.

From a decision tree, patterns, values, and conditions of independent attributes that make up a target class can be extracted. The overall quality of the tree is measured by *success rate*, while the quality of each class is measured by *true positive rate*, *false positive rate*, and *precision*. They are computed from confusion matrix (Figure 2.6) as follows :

$$\text{success rate} = \{\sum_{i=1}^{n}\} \,/\, \text{total} \quad (1)$$

$$\text{TP(a)} = f_{aa} \,/\, actual(a) \quad (2)$$

$$\text{FP(a)} = (predict(a) - f_{aa}) \,/\, (total - actual(a)) \quad (3)$$

$$\text{precision(a)} = f_{aa} \,/\, predict(a) \quad (4)$$

*TP(a)*, denoting true positive rate of class a, indicates the accuracy of predicting that data instances belong to class *a*. *FP(a)*, denoting false positive rate of class *a*, indicates the classifier's bias toward class *a*, which leads to false alarms. *Precision(a)* indicates the consistency of predicting that the instances belong to class *a*. Classes with high *TP(a)*, high *precision(a)*, but low *FP(a)* are preferred. To prevent overfitting, parameters (such as minimum node frequency) will be adjusted when pruning the trees.

Class predicted by classifier

| Actual Class | *a* | *b* | *c* | |
|---|---|---|---|---|
| *a* | *faa* | *fab* | *fac* | $\sum_{k=1}^{n} f_{ak}$ |
| *b* | *fba* | *fbb* | *fbc* | . |
| *c* | *fca* | *fcb* | *fcc* | . |
| | $\sum_{i=1}^{n} f_{ia}$ | . | . | total |

*fac* = Frequency of instances whose actual class is *a* and predicted class is *c*

*total* = Total instances in the dataset

$\sum_{k=1}^{n} f_{ak}$ = *actual(a)* or frequency of instances whose actual class is *a*

$\sum_{i=1}^{n} f_{ia}$ = *predict(a)* or frequency of instances whose predicted class is *a*

**Figure 2.6** Confusion Matrix, with classes *a*, *b*, and *c* [23]

## 2.8 Related Research

Wiyada Nintranon, 2007, studied in topic the development of the Life Distress Inventory computerized version. The object of this study was to develop the Life Distress Inventory (LDI) computerize version. Internal consistency reliability, correlation and satisfaction of the LDI computerized were examined. The subjects testing and evaluating the LDI computerized version in this study were 140 studying at Matthayom 6 (grade 12) of Watraikhingwittaya, Nakornprathom. The result found that the LDI computerized version has high internal consistency reliability at 0.916. The LDI computerized version has a positive relationship with the LDI traditional version at a high correlation and with significant difference at $p<0.01$ ($r=0.879$). Regarding the level of satisfaction, it was found that the subjects were very satisfied with the LDI computerized version. Furthermore, the subjects preferred the computerized version to the traditional version.

In conclusion, the results in this study indicate that the LDI computerized version is an efficient psychological health instrument that can be used as a practical alternative to the traditional version. Further investigation should be made regarding the efficiency of the instrument in other population [24].

Shannon L.Currie, Patrick J.,McGrath and Victor Day studied about development and usability of an online CBT program for symptoms of moderate depression, anxiety, and stress in post-secondary student.This study describes the development and usability testing of a new cognitive behavioral therapy-based program, ‘‘Feeling Better” designed to reduce symptoms of emotional distress in post-secondary students. An iterative qualitative usability testing approach was used to assess the program’s usability (i.e., ease of navigation, clarity, efficiency and acceptability). Three cycles of participant feedback and feedback from counselling centre staff was coded and used to iteratively modify the interface. Changes were both structural (e.g., shortening sections)and stylistic (e.g., aesthetic features), remodeling ‘‘Feeling Better” into a user-friendly platform ready to be used and its effects evaluated in further studies.The result was found that Internet-based intervention programs can improve mental health outcomes, and may offer a novel medium for reducing emotional distress in post-secondary students[25].

Erik Hedman,et al.studied about internet administration of self-report measures commonly used in research on social anxiety disorder: A psychometric evaluation. The purpose of this study was to test the equivalence of paper-and-pencil and Internet administered versions of self-report questionnaires used in social phobia research. This study analyzed data from two trials in which samples were recruited in a similar manner. One sample (N = 64) completed the paper-and-pencil version of questionnaires and the second sample (N = 57) completed the same measures online. They included the Liebowitz Social Anxiety Scale-self-assessment (LSAS-SR), the Social Interaction and Anxiety Scale (SIAS), and the Social Phobia Scale (SPS) as measures of social anxiety. Also included were the Montgomery Asberg Depression Rating Scale-self-assessment (MADRS-S), the Beck Anxiety Inventory (BAI), and the Quality of Life Inventory (QOLI). Results showed equivalent psychometric properties across administration formats. Cronbach’s a ranged between 0.77 and 0.94. There was an indication of a somewhat higher construct validity when participants filled out questionnaires using paper-and-pencil. Conclude that the LSAS-SR, SIAS, and SPS can be administered via the Internet with maintained psychometric properties[26].

# CHAPTER III
# MATERIALS AND METHODS

This chapter explained the step of research methodology based on Software Development Life Cycle (SDLC)

## 3.1 Research Methodology

### 3.1.1 Gathering data and information

Investigate each function to develop Online Rating Scale for The Life Distress Inventory. In this step, to research documents, publications, articles, journals and books are collected from secondary sources. This information is about the data of The Life Distress Inventory that questionnaire about situation causing stress in daily life, web-based application development. Also with gathering by professional research to define and requirement analysis of existing testing system.All of this information will be based on knowledge for develop Online Rating Scale for The Life Distress Inventory.

### 3.1.2 System analysis

This step is to review the existing data for analysis the system for development web-based application of The Life Distress Inventory. The details works as the following:

1) Analyze the data which collected in data collection step and classify it into the group with according of questionnaire to measurement situation causing stress in daily life database.

2) Analyze the system. The components of structural analysis are the following:

2.1 Data flow diagram (DFD) for database design will be created and added references to the ways in which the data stores, data flow, and processes will be implemented.

2.2 Entity relationship diagram (ERD) contains references how data will be stored in a file of database table.

**3.1.3 System Design**

1) Design the database. The detailed works with the following:

- Database design will use ERD from analysis phase to present designing details, and metadata are including describing the data model components.

- Database security will be design as authorized users can manipulate the data by two ways; manipulate through the DBMS and web database application; security database system requires the user name and password for each user can get into the database system. The general user can test on questionnaire of The Life Distress Inventory on the Internet.

2) Designing website, webpage structure and navigation.

3) Design the user interface that consists of interface design, controlling data access design, dialogue design. The users interface design is process of defining how the users will interact with the system and the nature of the inputs and outputs that the system accepts and produces.

4) Building a prototype for presenting and discussion.

5) Adjusting the design until it meets the requirement of user.

6) Building data definition in Data Dictionary.

7) Building program specifications by mapping class diagram to source code.

8) Selecting software, applications, and hardware that the system will require.

**3.1.4 System development**

After analyzing and designing step for the database and process of web-based application is completed, then developing this system with software applications.

### 3.1.5 System Implementation and Evaluation

After develop is completed. It is necessary to test it for finding errors. The debugging process will be used if there is error occurred. The system will be simulated and run on web server program with a personal computer, and users will answer a questionnaire for evaluate the system.

## 3.2 Materials

### 3.2.1 Hardware Specification (Minimum requirement)

**Server**

- CPU : Pentium IV 2.0 GHz.
- Memory : 1 GB.
- Hard Disk : 80 GB.
- Monitor : VGA
- Peripheral : Mouse and keyboard
- Network Device : 10/100/1000 Ethernet NIC

**Client**

- CPU : Pentium IV 1.6 GHz.
- Memory : 512 MB
- Hard Disk : 40 GB.
- Monitor : VGA
- Peripheral : Mouse and keyboard
- Network Device : 10/100/1000 Ethernet NIC and Internet connection for public access.

### 3.2.2 Software Specification

**Server**

- Operating System :Microsoft Windows Server 2008, released September 23$^{rd}$, 2008.
- Database Server : Microsoft SQL
- Web Server : Internet Information Services 6 (IIS 6)

**Client**

- Operating System : Microsoft Windows XP Service Pack3
- Web Browsers : Microsoft Internet Explorer 6.0

## 3.3 Gantt Chart for System Development Life Cycle

| ID | Steps of Methodology | 2009 | | | | | | 2010 | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | July | Aug | Sep | Oct | Nov | Dec | Jan | Feb | Mar | Apr | May | Jun | July | Aug | Sep | Oct | Nov | Dec |
| 1 | Preliminary Investigate | ■ | ■ | | | | | | | | | | | | | | | | |
| 2 | Gathering Data and Information | | | ■ | ■ | | | | | | | | | | | | | | |
| 3 | System Analysis | | | | | ■ | ■ | ■ | | | | | | | | | | | |
| 4 | System Design | | | | | | | | ■ | ■ | ■ | | | | | | | | |
| 5 | System Development | | | | | | | | | | ■ | ■ | ■ | | | | | | |
| 6 | Implement and Evaluation | | | | | | | | | | | | ■ | ■ | ■ | ■ | | | |
| 7 | Documentation | | | | | | | | | | | | | | | | ■ | ■ | ■ |

# CHAPTER IV
# RESULTS

The result of this study is online web questionnaires database system for measure of the life distress inventory .This chapter presents and describes system analysis , system design, system development and the result of testing.

## 4.1 System Analysis

A development of online web questionnaires database system for measure of the life distress inventory consists of three parts: registration, questionnaire, report result

### 4.1.1 Registration

Users have to register as a member before they use the system. This part collects user information such as username and password that use to verify user. And collects user personal information such as gender, education, job and age into the database.

### 4.1.2 Questionnaire

After users log in to the system. Users can answer questions about the life distress inventory which have 20 items.

### 4.1.3 Report Results

After users had answered the question admin can view the result. Result is reported in by group such as gender, education, income and range of age. The selection was based on what you want to see.

## 4.2 System Design

### 4.2.1 Data Flow Diagram

The data flow diagram of the online rating scale for The Life Distress Inventory consists of two levels. It shows functional dependency of the system including the identification of the input and output data.

**Context Diagram ( Level 0 )**

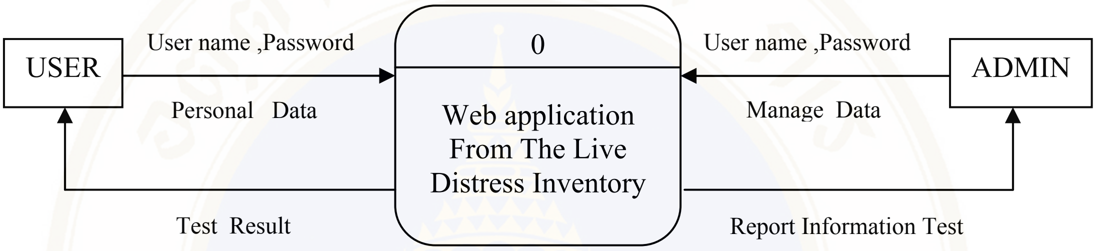



**Figure 4.1** Context Diagram of the online rating scale for The Life Distress Inventory

**Level 1**

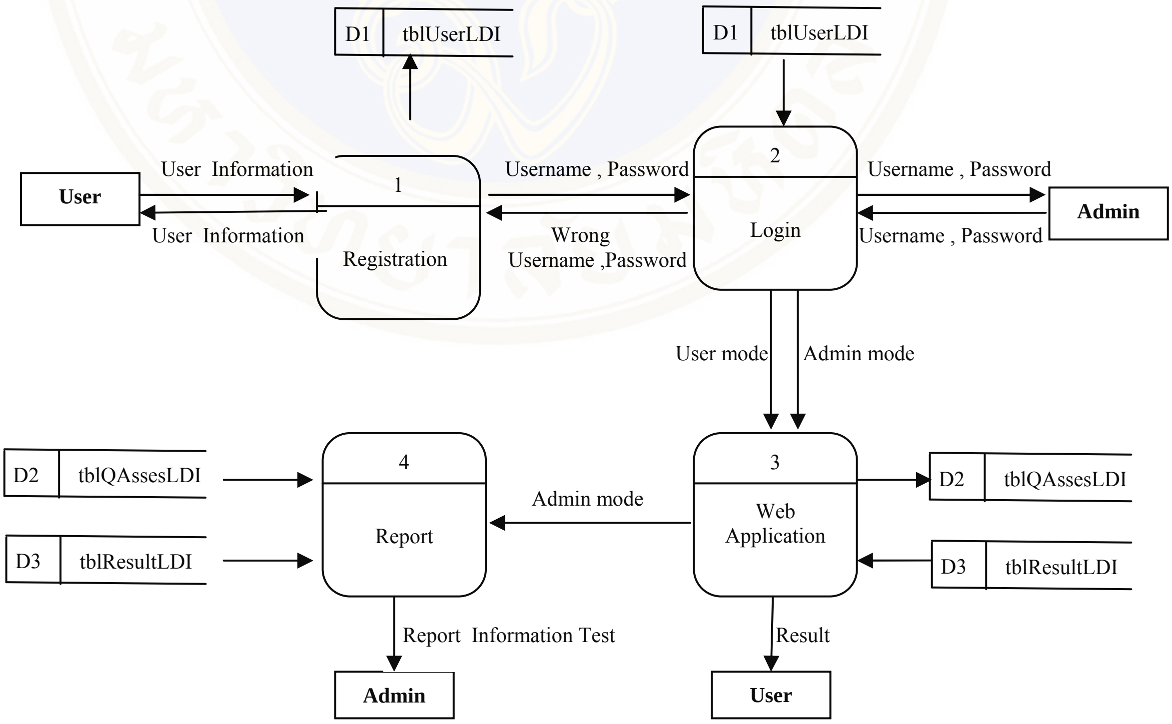



**Figure 4.2** Data flow diagram of the online rating scale for The Life Distress Inventory.

### 4.2.2 Database Design

#### 4.2.2.1 E-R Diagram

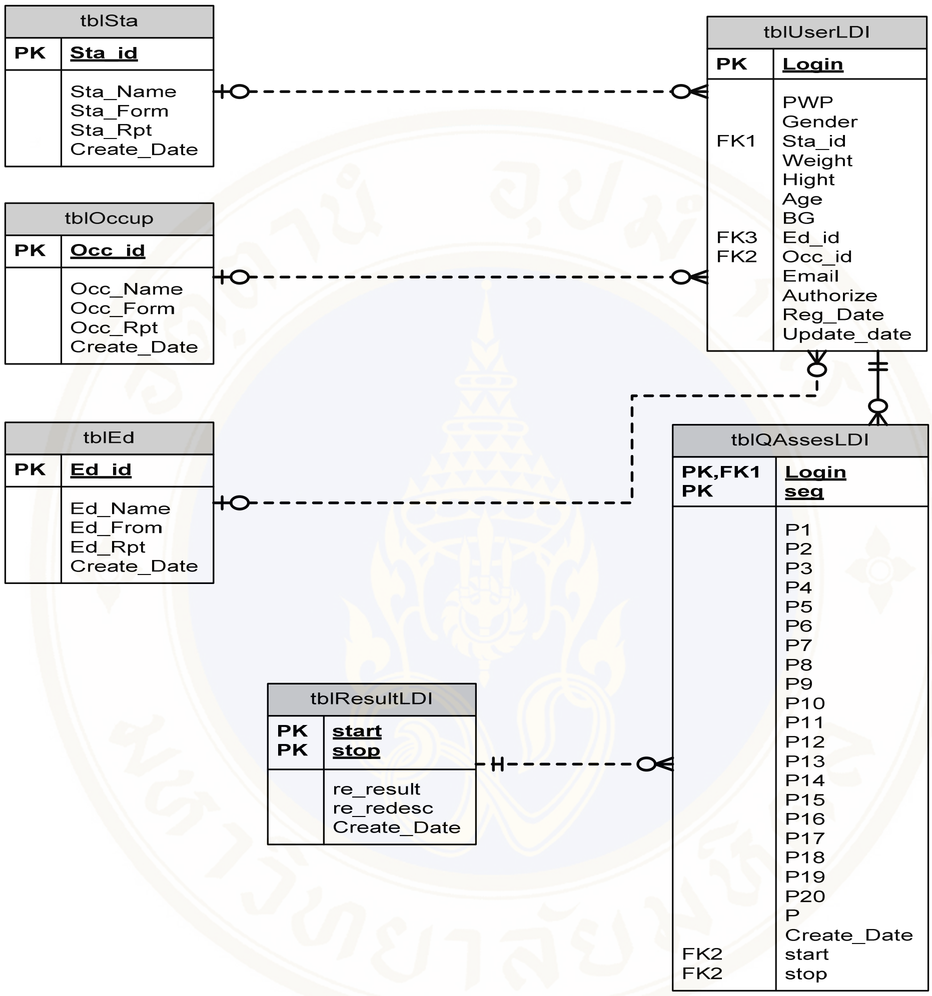



**Figure 4.3** Entity Relationship Diagram of the online rating scale for The Life Distress Inventory.

#### 4.2.2.2 Data Dictionary

**Table 4.1** Data dictionary of TblEd

| Field Name | Data Type | Length | Key | Description |
|---|---|---|---|---|
| Ed_id | smallint | | PK | Education ID |
| Ed_Name | nvarchar | 50 | | Education Name |
| Ed_From | nvarchar | 10 | | Education From |
| Ed_Rpt | nvarchar | 10 | | Educatiom Report |
| Create_Date | datetime | | | Create Date |

**Table 4.2** Data Dictionary of TblOccup

| Field Name | Data Type | Length | Key | Description |
|---|---|---|---|---|
| Occ_id | smallint | | PK | Occupation ID |
| Occ_Name | nvarchar | 50 | | Occupation Name |
| Occ_From | nvarchar | 10 | | Occupation From |
| Occ_Rpt | nvarchar | 10 | | Occupation Report |
| Create_Date | datetime | | | Create Date |

**Table 4.3** Data Dictionary of TblSta

| Field Name | Data Type | Length | Key | Description |
|---|---|---|---|---|
| Sta_id | smallint | | PK | Status ID |
| Sta_Name | nvarchar | 10 | | Status Name |
| Sta_From | nvarchar | 10 | | Status From |
| Sta_Rpt | nvarchar | 10 | | Status Report |
| Create_Date | datetime | | | Create Date |

**Table 4.4** Data Dictionary of TblUserLDI

| Field Name | Data Type | Length | Key | Description |
|---|---|---|---|---|
| Login | nvarchar | 50 | PK | User Name |
| PWD | nvarchar | 20 | | Password |
| Gender | nvarchar | 5 | | Gender |
| Sta_id | smallint | | FK | Status ID |
| Weight | smallint | | | Weight |
| Hight | smallint | | | Hight |
| Age | smallint | | | Age |
| BG | nvarchar | 10 | | Blood Group |
| Ed_id | smallint | | FK | Education ID |
| Occ_id | smallint | | FK | Occupation ID |
| Email | nvarchar | 50 | | Email Address |
| Authorize | nvarchar | 10 | | Types of users. |
| Reg_Date | datetime | | | Registration Date |
| Update_Date | datetime | | | Update Date |

**Table 4.5** Data Dictionary of TblResultLDI

| Field Name | Data Type | Length | Key | Description |
|---|---|---|---|---|
| start | int | | PK | Starting score |
| stop | int | | PK | Final score |
| side_id | nvarchar | 10 | PK/FK | Side ID |
| re_result | int | | | Test Results |
| re_redesc | int | | | Description of the test results |
| Create_Date | datetime | | | Create Date |

**Table 4.6** Data Dictionary of TblQAssesLDI

| Field Name | Data Type | Length | Key | Description |
|---|---|---|---|---|
| Login | nvarchar | 50 | PK | Login |
| Seq | smallint | | PK | Sequence |
| P1 | smallint | | | Question 1 |
| P2 | smallint | | | Question 2 |
| P3 | smallint | | | Question 3 |
| P4 | smallint | | | Question 4 |
| P5 | smallint | | | Question 5 |
| P6 | smallint | | | Question 6 |
| P7 | smallint | | | Question 7 |
| P8 | smallint | | | Question 8 |
| P9 | smallint | | | Question 9 |
| P10 | smallint | | | Question 10 |
| P11 | smallint | | | Question 11 |
| P12 | smallint | | | Question 12 |
| P13 | smallint | | | Question 13 |
| P14 | smallint | | | Question 14 |
| P15 | smallint | | | Question 15 |
| P16 | smallint | | | Question 16 |
| P17 | smallint | | | Question 17 |
| P18 | smallint | | | Question 18 |
| P19 | smallint | | | Question 19 |
| P20 | smallint | | | Question 20 |
| P | smallint | | | Total score |
| Create_Date | datetime | | | Create Date |

### 4.2.3 User Interface Design

#### 4.2.3.1 Interface Design

The Graphic User Interface (GUI) input control is used for helping user to use this web database easier. The GUI input controls that are used in this web database are text box, radio button, check box and drop down list box.

#### 4.2.3.2 Controlling Data Access Design

Users will have to login with their username and password before access the web application. The level of authorization to the web application is divided into 2 groups : user and administrator. Administrator will be the one who can access to database manipulation and also can access to an parts of the web application, while user can not.

After users have to register their details and information to get username and password. It should be required login ID and password for individual; consequently, if users sign up as member and login, users will be able to test the questionnaire by following each step of questions.

## 4.3. System Development

**Online Rating Scale for The Life Distress Inventory**

This system can be divided into 4 parts as follow

### 4.3.1 Register

When user access to system. System responses with first page as figure 4.4. If user still not register in system, user would click "Register New User".

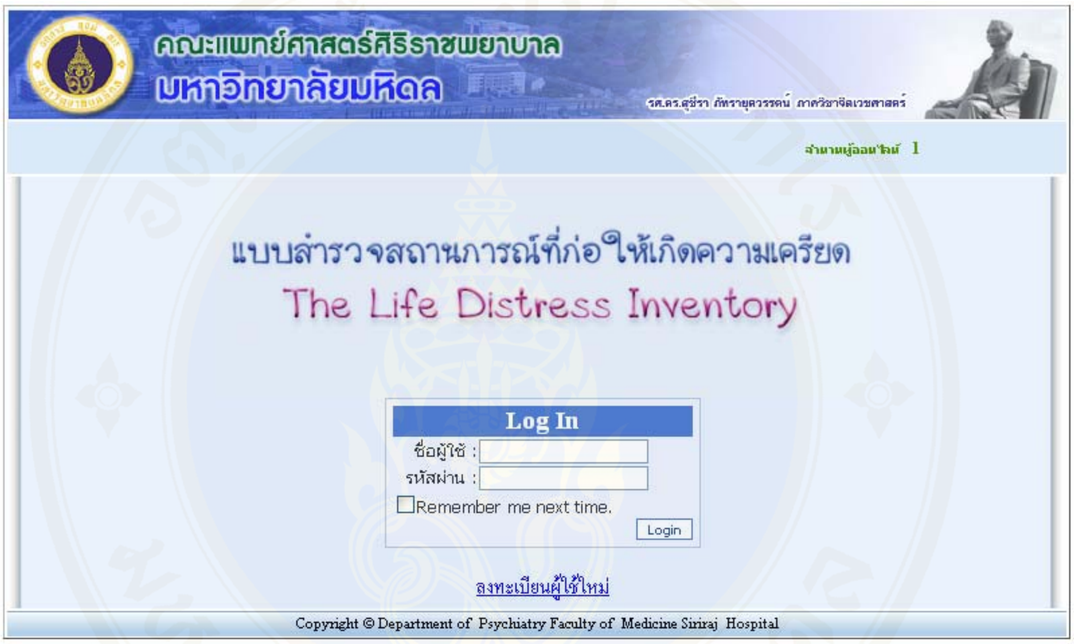



**Figure 4.4** The screen of home page of the online rating scale for The Life Distress Inventory for process registration and sign in.

The main function of registration page is filling full the personal details to the dialog box. The dialogs contain the details of:

- User Name
- Password
- Confirm password
- Gender
- Status
- Weight
- Height
- Age

- Blood Group
- Education
- Occupation
- Email address

All registration must be complete. The system Admin will be authorized the credible users by consider from the registration details. User must be filling the detail in every dialog. After filling the information, push the "SUBMIT" button to save information. If don't want to submit just push "CANCEL" button to cancel all the filling information. After the information is filled, all information will store in the database.Then, can register on system by filling in username and password in figure 4.4


**Figure 4.5** The screen of home page of the online rating scale for The Life Distress Inventory for process of participant registration.

### 4.3.2 Questionnaire Administration

Users , already have username and password, can access in system by enter username and password in Figure 4.4. System will verify username and password with its database and return page as figure 4.6. Figure 4.6, the start page of system, The start page describe about the suggestion before start to do questionnaire about The Life Distress Inventory . The main function of start page will contain start, edit profile and logout.


**Figure 4.6** The screen of start page of the online rating scale for The Life Distress Inventory for process of start The Life Distress Inventory and edit personal profile.

After User select start button, User can start to do questionnaire about The Life Distress Inventory with all 20 item. Each questionnaire measuring events causing stress in daily life using five level ordinals rating scales contain most, much, medium, little and none.

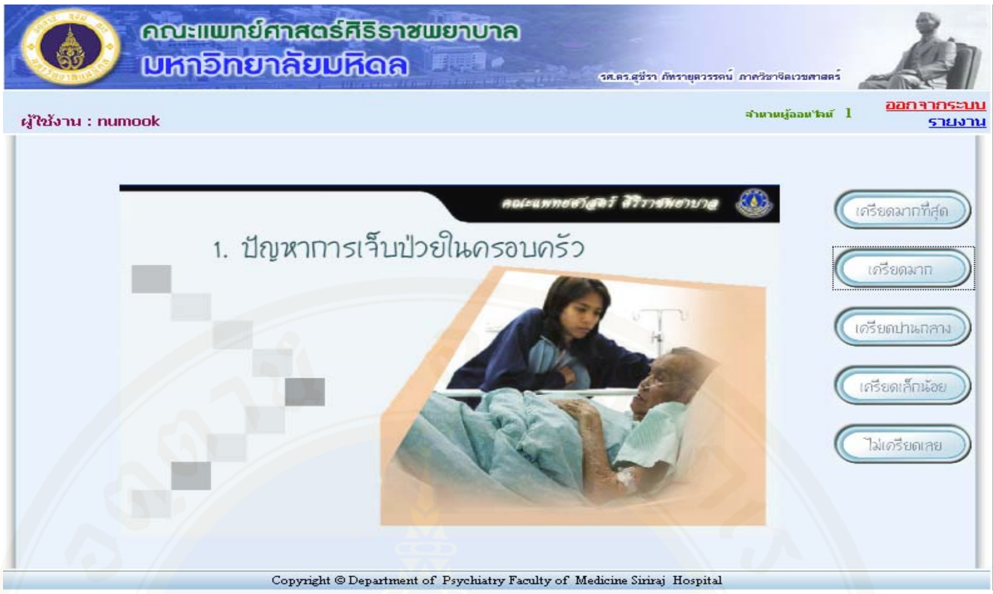



**Figure 4.7** The screen of the online rating scale for The Life Disress Inventory for process of questionnaire testing.

### 4.3.3 Results

When user did the questionnaires to the last question, the results page appear and describe in wording about stress level of user have in events causing stress in daily life after finished testing of questionnaires.


**Figure 4.8** The online rating scale for The Life Distress Inventory provide report the results for user.

### 4.3.4 Questionnaires of user's satisfaction

After users finished on their The Life Distress Inventory's testing they will comment on the questionnaires of user satisfaction. The satisfaction of questionnaires will be comment by users who have test on the online rating scale for The Life Distress Inventory .

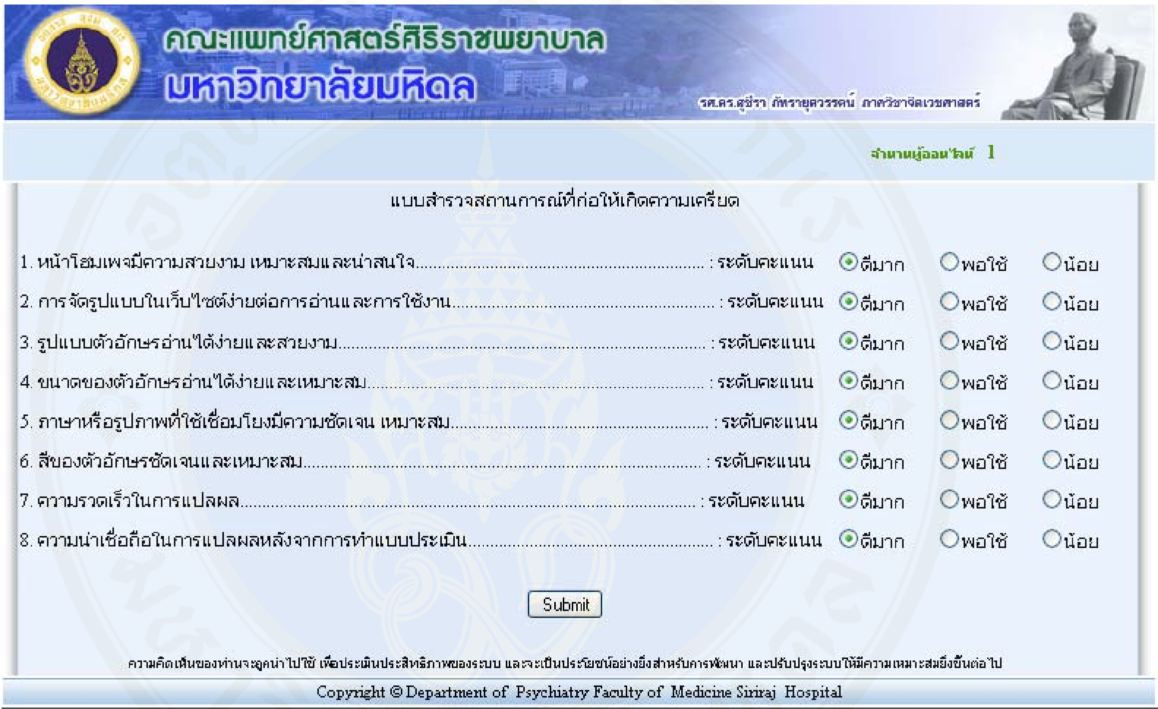



**Figure 4.9** The screen of satisfaction of users of the online rating scale for The Life Distress Inventory for the questionnaires of the Web-page.

## 4.4. Implementation and testing

This application use Microsoft SQL as database. ASP.NET is server side programming. After completely coding this application then we will test this to ensure that it work properly. Testing is divided into two parts:

### 4.4.1 Unit testing

We will focus on each unit of data retrieval system and data management system. The users interacted with system through the developed forms so that the function of each command button on each page was tested.

#### 4.4.2 System Testing

After all of units were tested, the system testing would be done to ensure that the system work properly. The information retrieval system might fail to work together, the researcher test the system. And this system will be tested by the researchers and specialists at the Department of psychiatry, Faculty of Medicine, Siriraj Hospital, Mahidol University.

#### 4.4.3 System Evaluation

The application for web questionnaire The Life Distress Inventory browser had URL: http://119.46.192.18/LDI. The implementation was experimented for 300 Record. The evaluation of the system was evaluated by the researchers and specialists at the Department of psychiatry, Faculty of Medicine, Siriraj Hospital, Mahidol University. The evaluation process is collecting user personal data who was answer the questionnaires of The Life Distress Inventory. The questionnaire consists 20 items. The result was retrieved after finished answer the questionnaires for users. The report will be generated for the researchers, specialists and doctors.

### 4.5 Clustering's results.

#### 4.5.1 The Summary of Participant Information

After gathered data from participant within 3 months, since 1 May 2010 until 31 July 2010, data specifications summarized from 308 participants were shown in Table.4.7

**Table 4.7** Summary of participant information (N = 308)

| Characteristic of participant | N | % |
|---|---|---|
| **1. Gender** | | |
| • Male | 84 | 27.27 |
| • Female | 224 | 72.73 |

**Table 4.7** Summary of participant information (continued)

| Characteristic of participant | N | % |
|---|---|---|
| **2. Status** | | |
| • Single | 248 | 80.52 |
| • Married | 51 | 16.56 |
| • Widowed | 5 | 1.62 |
| • Divorced | 4 | 1.30 |
| **3. Age** (years ) | | |
| • 19 - 25 | 109 | 35.39 |
| • 26 – 45 | 174 | 56.49 |
| • 46 – 60 | 25 | 8.12 |
| **4. Education** | | |
| • High School | 7 | 2.27 |
| • Diploma | 7 | 2.27 |
| • Bachelor Degree | 227 | 73.70 |
| • Master Degree | 58 | 18.83 |
| • Ph.D. | 9 | 2.93 |
| **5. Occupation** | | |
| • unemployed | 8 | 2.60 |
| • Students | 61 | 19.81 |
| • Employees | 59 | 19.15 |
| • Government service | 150 | 48.70 |
| • Freelance | 21 | 6.81 |
| • Private business | 9 | 2.93 |

### 4.5.2 WEKA's Results.

Weka, a collection of machine learning algorithms for solving data mining problems, was developed at the University of Waikato in New Zealand. "Weka"stands for the Waikato Environment for Knowledge Analysis. Weka is open source software issued under the GNU public license[27]. Weka provides implementations of learning algorithms that one can easily apply to datasets. It also includes a variety of tools for transforming datasets. The workbench includes methods for all the standard data mining problems: regression, classification, clustering, association rule mining, and attribute selection[23].

In this study was used *SimpleKMeans* for construction the clustering. *SimpleKMeans* was used as the main algorithm in clustering experiment because it is simple and its results are easy to interpret[23].

For the analysis of the life distress inventory data, clustering is first performed.Typically, tuples with similar attribute values are clustered into the same class.

Once the classes are derived, a classification algorithm can be applied in order to generate classification rules for each class. The Weka's module, *J48*, was employed for the construction the C4.5 decision trees. The overall quality of the tree is measured by *success rate*, while the quality of each class is measured by *true positive rate*, *false positive rate* and *precision*. They are computed from confusion matrix (Figure 4.10) as follows [23]:

$$\text{success rate} = \{\sum_{i=1}^{n}\} / \text{total} \quad (1)$$

$$TP(a) = faa / actual(a) \quad (2)$$

$$FP(a) = (predict(a) - faa) / (total - actual(a)) \quad (3)$$

$$precision(a) = faa / predict(a) \quad (4)$$

*TP(a)*, denoting true positive rate of class a, indicates the accuracy of predicting that data instances belong to class *a*. *FP(a)*, denoting false positive rate of class *a*, indicates the classifier's bias toward class *a*, which leads to false alarms. *Precision(a)* indicates the consistency of predicting that the instances belong to class *a*.Classes with high *TP(a)*, high *precision(a)*, but low *FP(a)* are preferred.

Class predicted by classifier

| Actual Class | a | b | c | |
|---|---|---|---|---|
| a | *faa* | *fab* | *fac* | $\sum_{k=1}^{n} f_{ak}$ |
| b | *fba* | *fbb* | *fbc* | . |
| c | *fca* | *fcb* | *fcc* | . |
| | $\sum_{i=1}^{n} f_{ia}$ | . | . | total |

*fac* = Frequency of instances whose actual class is *a* and predicted class is *c*

*total* = Total instances in the dataset

$\sum_{k=1}^{n} f_{ak}$ = *actual(a)* or frequency of instances whose actual class is *a*

$\sum_{i=1}^{n} f_{ia}$ = *predict(a)* or frequency of instances whose predicted class is *a*

**Figure 4.10** Confusion Matrix, with classes *a*, *b*, and *c* [23]

In this study,the clustering experiment is divided into two experiments .

**4.5.2.1 Clustering experiment 1**

The algorithm is applied to each dataset seven times, with the number of cluster being 2, 3, 4, 5, 6, 7 and 8 (numClusters = 2, 3, 4, 5, 6, 7 or 8) – this parameter was adjusted during the experiment, until an optimal result was obtained.

The dataset was grouped into a predefined number of clusters (numCluster =2, 3, 4, 5, 6, 7 or 8). Then, cluster labels were attached to data instances. The *J48* classification was next performed, with cluster label being the target attribute. Summarized results are shows in Table 4.8.

**Table 4.8** Preliminary result of the Life Distress Inventory dataset *(target attribute = cluster)*

| Life Distress Inventory Dataset | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| numCluster | Success | Cluster 1 | | | Cluster 2 | | | Cluster 3 | | | Cluster 4 | | |
| | | TP rate | FP rate | Precision | TP rate | FP rate | Precision | TP rate | FP rate | Precision | TP rate | FP rate | Precision |
| 2 | 1.00 | 1.00 | 0 | 1.00 | 1.00 | 0 | 1.00 | - | - | - | - | - | - |
| 3 | 0.981 | 0.987 | 0.013 | 0.987 | 0.976 | 0.018 | 0.953 | 0.969 | 0 | 1.00 | - | - | - |
| 4 | 0.974 | 0.987 | 0.007 | 0.994 | 0.948 | 0.017 | 0.948 | 0.953 | 0.011 | 0.932 | 1.00 | 0 | 1.00 |
| 5 | 0.967 | 1.00 | 0.02 | 0.981 | 0.96 | 0.009 | 0.973 | 0.953 | 0.019 | 0.891 | 1.00 | 0 | 1.00 |
| 6 | 0.971 | 0.968 | 0.02 | 0.923 | 1.00 | 0.004 | 0.983 | 0.984 | 0 | 1.00 | 1.00 | 0 | 1.00 |
| 7 | 0.938 | 0.911 | 0.016 | 0.927 | 0.949 | 0.012 | 0.949 | 0.977 | 0.034 | 0.827 | 1.00 | 0.007 | 0.931 |
| 8 | 0.945 | 0.927 | 0.02 | 0.911 | 0.966 | 0.008 | 0.966 | 0.977 | 0.019 | 0.896 | 1 | 0.018 | 0.828 |

**Table 4.8** (continued) Preliminary result of the Life Distress Inventory dataset *(target attribute = cluster)*

| Life Distress Inventory Dataset | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| numCluster | Success | Cluster 5 | | | Cluster 6 | | | Cluster 7 | | | Cluster 8 | | |
| | | TP rate | FP rate | Precision | TP rate | FP rate | Precision | TP rate | FP rate | Precision | TP rate | FP rate | Precision |
| 2 | 1.00 | - | - | - | - | - | - | - | - | - | - | - | - |
| 3 | 0.981 | - | - | - | - | - | - | - | - | - | - | - | - |
| 4 | 0.974 | - | - | - | - | - | - | - | - | - | - | - | - |
| 5 | 0.967 | 0 | 0 | 0 | - | - | - | - | - | - | - | - | - |
| 6 | 0.971 | 0 | 0 | 0 | 1.00 | 0.014 | 0.968 | - | - | - | - | - | - |
| 7 | 0.938 | 0 | 0 | 0 | 1.00 | 0.005 | 0.991 | 0.50 | 0 | 1.00 | - | - | - |
| 8 | 0.945 | 0 | 0 | 0 | 1 | 0 | 1 | 0.833 | 0 | 1 | 0.556 | 0 | 1 |

When numCluster was set to 2-8, Cluster cannot separate participant who have status married.When numClusters were set to 5 , 6 , 7 and 8, *Cluster 5* had low *TP rates* . When numCluster was set to 2, ideal *Success*, *TP rates*, *FP rates* and *Precisions* were achieved, but it seemed over-optimistic. Therefore, numCluster = 3 and numCluster = 4 were optimal but numCluster = 3 yielded higher *Success* than numCluster = 4. Therefore, numCluster = 3 was chosen.

The cluster's result of the life distress inventory dataset in clustering experiment 1 is described as follow:

Number of iterations: 3

Within cluster sum of squared errors: 514.0

Missing values globally replaced with mean/mode

Cluster centroids:

| | | Cluster | | |
|---|---|---|---|---|
| Attribute | Full Data | 0 | 1 | 2 |
| | (308) | (162) | (75) | (71) |
| Gender | Female | Female | Female | Male |
| Status | Single | Single | Single | Single |
| Age | 26-45 | 26-45 | 19-25 | 26-45 |
| Education | Bachelor | Bachelor | Bachelor | Bachelor |
| Occupation | Government | Government | Students | Government |
| Output | Little stress | Little stress | Moderate stress | Little stress |

Clustered Instances

| | |
|---|---|
| 0 | 162 ( 53%) |
| 1 | 75 ( 24%) |
| 2 | 71(23%) |

The cluster description are as follows:

Cluster 0 : Gender female, Status single ,Age 26-45,Education Bachelor's degree and occupation government **then** they will have Little Stress ( 162 instances ).

Cluster 1 : Gender female, Status single ,Age 19-25,Education Bachelor's degree and occupation students **then** they will have Moderate stress ( 75 instances ).

Cluster 2 : Gender male, Status single ,Age 26-45,Education Bachelor's degree and occupation government **then** they will have Little Stress ( 71 instances ).

**The classification profile of each cluster (numCluster = 3)**

Table 4.9 summarizes results and classification metrics of decision tree, with *cluster* being the target attribute. There are 11 leaf nodes in total. The findings can be summarized as follows :

**Table 4.9** Preliminary Result (target attribute = cluster)

| **The life distress inventory dataset** | | | |
|---|---|---|---|
| **Overall Stat.** | **Class 1** | **Class 2** | **Class 3** |
| Leaf nodes = 11<br>Success = 0.981 | TP rate = 0.987<br>FP rate = 0.013<br>Precision = 0.987 | TP rate = 0.976<br>FP rate = 0.018<br>Precision = 0.953 | TP rate = 0.969<br>FP rate = 0<br>Precision = 1 |

**Class 1:** Cluster 0

a. Members of this group are female, age 19-25 years old and have a career as a government service.

b. Members of this group are female, age 26-45 years old.

c. Members of this group are female, age 46-60 years old.

**Class 2:** Cluster 1

a. Members of this group are female, age between 19-25 years and unemployed.

b. Members of this group are students, gender female and age between 19-25 years.

c. Members of this group are female, age between 19-25 years and have a career as employees.

d. Members of this group are female, age between 19-25 years and have a career as freelance.

e. Members of this group are female, age between 19-25 years and have a career as a private business.

f. Members of this group are male, age between 19-25 years .

**Class 3:** Cluster 2

a. Members of this group are male, age between 26-45 years .

b. Members of this group are male, age between 46-60 years .

**4.5.2.2 Clustering experiment 2**

The algorithm is applied to each dataset four times, with the number of cluster being 9, 10, 11 and 12 (numClusters = 9, 10, 11 or 12) – this parameter was adjusted during the experiment, until an optimal result was obtained.

The dataset was grouped into a predefined number of clusters (numCluster =9, 10, 11 or 12). Then, cluster labels were attached to data instances. The *J48* classification was next performed, with cluster label being the target attribute. Summarized results are shows in Table 4.10.

**Table 4.10** Preliminary result of the Life Distress Inventory dataset *(target attribute = cluster)*

| Life Distress Inventory Dataset | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| numCluster | Success | Cluster 1 | | | Cluster 2 | | | Cluster 3 | | | Cluster 4 | | |
| | | TP rate | FP rate | Precision | TP rate | FP rate | Precision | TP rate | FP rate | Precision | TP rate | FP rate | Precision |
| 9 | 0.951 | 0.92 | 0.012 | 0.939 | 1 | 0.004 | 0.983 | 1 | 0.019 | 0.898 | 1 | 0.01 | 0.864 |
| 10 | 0.938 | 0.96 | 0.023 | 0.889 | 1 | 0.008 | 0.967 | 0.977 | 0.011 | 0.935 | 1 | 0.017 | 0.792 |
| 11 | 0.948 | 0.94 | 0.016 | 0.922 | 0.983 | 0.004 | 0.983 | 0.984 | 0.012 | 0.952 | 1 | 0.01 | 0.864 |
| 12 | 0.916 | 0.98 | 0.016 | 0.925 | 1 | 0.004 | 0.983 | 0.885 | 0.027 | 0.868 | 1 | 0.003 | 0.917 |

**Table 4.10** (continued) Preliminary result of the Life Distress Inventory dataset *(target attribute = cluster)*

| Life Distress Inventory Dataset | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| numCluster | Success | Cluster 5 | | | Cluster 6 | | | Cluster 7 | | | Cluster 8 | | |
| | | TP rate | FP rate | Precision | TP rate | FP rate | Precision | TP rate | FP rate | Precision | TP rate | FP rate | Precision |
| 9 | 0.951 | 0 | 0 | 0 | 1 | 0 | 1 | 0.667 | 0 | 1 | 0.556 | 0.003 | 0.833 |
| 10 | 0.938 | 0 | 0.003 | 0 | 1 | 0.005 | 0.99 | 0.538 | 0 | 1 | 0.444 | 0 | 1 |
| 11 | 0.948 | 0.75 | 0.013 | 0.6 | 0.975 | 0 | 1 | 0.5 | 0 | 1 | 0.556 | 0 | 1 |
| 12 | 0.916 | 0.875 | 0.003 | 0.875 | 0.975 | 0.004 | 0.987 | 0.25 | 0 | 1 | 0.444 | 0 | 1 |

**Table 4.10** (continued) Preliminary result of the Life Distress Inventory dataset *(target attribute = cluster)*

| Life Distress Inventory Dataset | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| numCluster | Success | Cluster 9 | | | Cluster 10 | | | Cluster 11 | | | Cluster 12 | | |
| | | TP rate | FP rate | Precision | TP rate | FP rate | Precision | TP rate | FP rate | Precision | TP rate | FP rate | Precision |
| 9 | 0.951 | 1 | 0.007 | 0.846 | - | - | - | - | - | - | - | - | - |
| 10 | 0.938 | 0.667 | 0.003 | 0.8 | 0.8 | 0 | 1 | - | - | - | - | - | - |
| 11 | 0.948 | 0.857 | 0.003 | 0.857 | 1 | 0 | 1 | 1 | 0 | 1 | - | - | - |
| 12 | 0.916 | 0.857 | 0.003 | 0.857 | 1 | 0 | 1 | 1 | 0.003 | 0.857 | 0.588 | 0.031 | 0.526 |

When numCluster was set to 9-12, Cluster can separate participant who have status married. When numCluster = 9 and numCluster = 11 were optimal but numCluster = 9 yielded higher *Success* than numCluster = 11. Therefore, numCluster = 9 was chosen.

The cluster's result of the Life distress Inventory dataset in clustering experiment 2 is described as follow:

Number of iterations: 3

Within cluster sum of squared errors: 406.0

Missing values globally replaced with mean/mode

Cluster centroids:

| | | Cluster | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Attribute | Full Data | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| | (308) | (84) | (39) | (35) | (27) | (23) | (4) | (71) | (20) | (5) |
| Gender | Female | Female | Female | Male | Male | Female | Female | Female | Male | Male |
| Status | Single | Single | Single | Single | Single | Single | Single | Single | Single | Married |
| Age | 26-45 | 19-25 | 19-25 | 26-45 | 26-45 | 19-25 | 19-25 | 26-45 | 26-45 | 19-25 |
| Education | Bachelor | Bachelor | Bachelor | Bachelor | Bachelor | Bachelor | Diploma | Bachelor | Bachelor | Bachelor |
| Occupation | Government | Government | Students | Freelance | Government | Students | Government | Government | Employees | Employees |
| Output | Little stress | Little stress | Little stress | Little stress | Normal | Moderate stress | Little stress | Normal | Moderate stress | Severe stress |

Clustered Instances

| | |
|---|---|
| 0 | 84 ( 27%) |
| 1 | 39 ( 13%) |
| 2 | 35 ( 11%) |
| 3 | 27 ( 9%) |
| 4 | 23 ( 7%) |

Clustered Instances

| | |
|---|---|
| 5 | 4 ( 1%) |
| 6 | 71 ( 23%) |
| 7 | 20 ( 6%) |
| 8 | 5 ( 2%) |

The cluster description are as follows:

Cluster 0 : Gender female, Status single ,Age 19-25,Education Bachelor's degree and occupation government **then** they will have Little Stress ( 84 instances ).

Cluster 1 : Gender female, Status single ,Age 19-25,Education Bachelor's degree and occupation students **then** they will have Little Stress ( 39 instances ).

Cluster 2 : Gender male, Status single ,Age 26-45,Education Bachelor's degree and occupation freelance **then** they will have Little Stress ( 35 instances ).

Cluster 3 : Gender male, Status single ,Age 26-45,Education Bachelor's degree and occupation government **then** they will have Normal Stress ( 27 instances ).

Cluster 4 : Gender female, Status single ,Age 19-25,Education Bachelor's degree and occupation students **then** they will have Moderate Stress ( 23 instances ).

Cluster 5 : Gender female, Status single ,Age 19-25,Education Diploma and occupation government **then** they will have Little Stress ( 4 instances ).

Cluster 6 : Gender female, Status single ,Age 26-45,Education Bachelor's degree and occupation government **then** they will have Normal Stress ( 71 instances ).

Cluster 7 : Gender male, Status single ,Age 26-45,Education Bachelor's degree and occupation employees **then** they will have Moderate Stress ( 20 instances ).

Cluster 8 : Gender male, Status married ,Age 19-25, Education Bachelor's degree and occupation employees **then** they will have Severe Stress (5 instances )

**The classification profile of each cluster (numCluster = 9)**

Table 4.11 summarizes results and classification metrics of decision tree, with *cluster* being the target attribute. There are 33 leaf nodes in total. The findings can be summarized as follows :

**Table 4.11** Preliminary Result (target attribute = cluster)

| The life distress inventory dataset | | | |
|---|---|---|---|
| **Overall Stat.** | **Class 1** | **Class 2** | **Class 3** |
| Leaf nodes = 33<br>Success = 0.951 | TP rate = 0.92<br>FP rate = 0.012<br>Precision = 0.939 | TP rate = 1<br>FP rate = 0.004<br>Precision = 0.983 | TP rate = 1<br>FP rate = 0.019<br>Precision = 0.898 |
| | **Class 4** | **Class 5** | **Class 6** |
| | TP rate = 1<br>FP rate = 0.01<br>Precision = 0.864 | TP rate = 0<br>FP rate = 0<br>Precision = 0 | TP rate = 1<br>FP rate = 0<br>Precision = 1 |
| | **Class 7** | **Class 8** | **Class 9** |
| | TP rate = 0.667<br>FP rate = 0<br>Precision = 1 | TP rate = 0.556<br>FP rate = 0.003<br>Precision = 0.833 | TP rate = 1<br>FP rate = 0.007<br>Precision = 0.846 |

**Class 1:** Cluster 0

a. Members of this group ,age between 19-25 years and unemployed.

b. Members of this group ,age between 19-25 years ,have a career as employees and gender female.

c. Members of this group ,age between 19-25 years and have a career as government service.

d. Members of this group ,age between 19-25 years and have a career as private business.

e. Members of this group ,age between 46-60 years and status single.

f. Members of this group ,age between 46-60 years and status widowed.

g. Members of this group ,age between 46-60 years and status divorce.

**Class 2:** Cluster 1

a. Members of this group are students and age between 19-25 years.

b. Members of this group are students, gender female and age between 26-45 years.

**Class 3:** Cluster 2

a. Members of this group, age between 19-25 years and have a career as freelance.

b. Members of this group, age between 26-45 years, gender male and unemployed.

c. Members of this group are students, gender male and age between 26-45 years.

d. Members of this group, age between 26-45 years , gender male ,have a career as employees and status single.

e. Members of this group, age between 26-45 years , gender male , have a career as employees and status widowed.

f. Members of this group, age between 26-45 years , gender male ,have a career as employees and status divorce.

g. Members of this group, age between 26-45 years , gender male and have a career as freelance.

h. Members of this group, age between 26-45 years , gender male and have a career as private business.

i. Members of this group, age between 26-45 years , gender female and have a career as freelance.

**Class 4:** Cluster 3

a. Members of this group, age between 26-45 years , gender male ,have a career as government service.

.

**Class 5:** Cluster 4 : No member in this group.

**Class 6:** Cluster 5

a. Members of this group ,age between 26-45 years, gender female and unemployed.

b. Members of this group ,age between 26-45 years, gender female and have a career as employees.

c. Members of this group ,age between 26-45 years, gender female, have a career as government service and education diploma.

d. Members of this group ,age between 26-45 years, gender female, have a career as government service and education bachelor degree.

e. Members of this group ,age between 26-45 years, gender female, have a career as government service and education high school.

f. Members of this group ,age between 26-45 years, gender female, have a career as government service, education master degree and status single.

g. Members of this group ,age between 26-45 years, gender female, have a career as government service, education master degree and status widowed.

h. Members of this group ,age between 26-45 years, gender female, have a career as government service, education master degree and status divorce.

i. Members of this group ,age between 26-45 years, gender female, have a career as government service and education Ph.D..

j. Members of this group ,age between 26-45 years, gender female and have a career as private business.

**Class 7:** Cluster 6

a. Members of this group ,age between 19-25 years, have a career as employees and gender male.

b. Members of this group, age between 26-45 years , gender male ,have a career as employees and status married.

**Class 8:** Cluster 7

a. Members of this group ,age between 26-45 years, gender female, have a career as government service, education master degree and status married.

**Class 9:** Cluster 8

a. Members of this group ,age between 46-60 years and status married.

# CHAPTER V
# DISCUSSIONS

Research Objective was to develop an online Web questionnaires database system for measure of The Life Distress inventory and to discover some knowledge hidden in The Life Distress Inventory database by using data mining techniques (Clustering and Classification) to reach the objective.

## 5.1 The efficiency of online rating scale for The Life Distress Inventory.

Online Web questionnaires database system for measure of The Life Distress Inventory is an application for displaying questionnaire, collecting questionnaire responses and generate detail result of response for further analysis.

The system is also developed for user friendly interface and easily to understand.User has to register to be a member of the Web site. In this step application will store all of users' data into the database as a profile. And user can be changed the profile later by click on "Edit profile" menu. After that users have to login to the system and can answer questionnaire for measure of The Life Distress Inventory. All of questionnaire information is stored in database and researcher can retrieve necessary data for analysis.

The advantage of this system can show as follow:

- Online rating scale for The Life Distress Inventory is web-based architecture. Therefore users can access to the system anywhere and anytime via web browser.
- Scoring was measure and computed immediately after answer all the questionnaires finished.
- Interpretation was generated immediately after participants finished answer all questionnaires. The participant will know the results of interpretation.

- All of questionnaire information is stored in database, and data remain as long as server not broken.

## 5.2 The satisfaction of respondents.

The satisfaction of respondents about the online rating scale for The Life Distress Inventory version obtained for interviewing the respondents, the result presented in figure 5.1

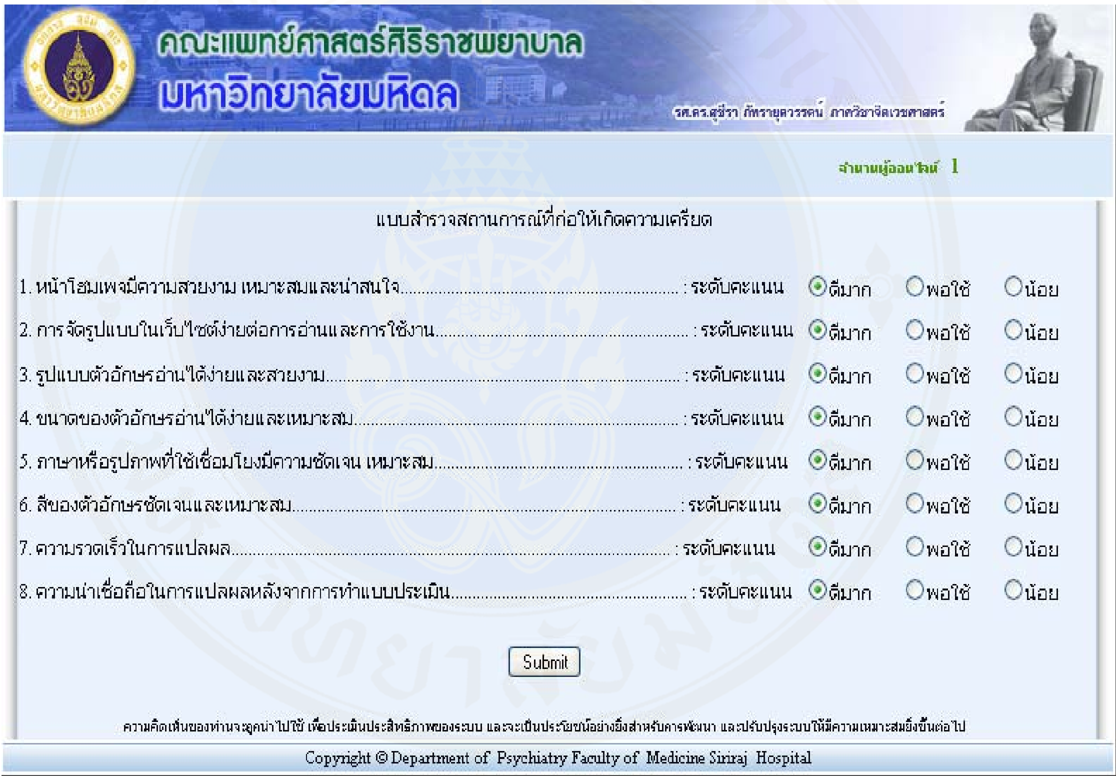



**Figure 5.1** The screen of satisfaction of users of the online rating scale for The Life Distress Inventory for the questionnaires of the Web-page.

The satisfaction of user was divided into eight contents format, appropriate of uses, interpretation and accuracy. The questionnaires of satisfaction users' comments in using online rating scale for The Life Distress Inventory. The score from user will make for many benefits in developments of online rating scale for The Life Distress Inventory.

The results of the number of satisfaction are mostly in the very good that can present about the range on the very satisfaction in all area, especially the satisfaction of the size of the text readable and appropriate that the range equal to 95.78%. Next satisfaction of users is speed of interpretation the range equal to 94.16%. Next satisfaction of users is the color fonts that the range equal to 93.51%. Next satisfaction of users is format the text readable and beautiful that the range equal to 93.51%. Next satisfaction of users is format web site easy to read and use that the range equal to 88.64%. Next satisfaction of users in area of language and the use of picture illustrations that the range equal to 87.34%. Next satisfaction of users is reliability in interpreting the results after finished the questionnaires of online rating scale for the life distress inventory that make the range equal to 82.79%.The last next score rate will be the beautiful and interest of online rating scale for The Life Distress Inventory web page that the range equal to 82.14%.That's the range of the score can present that most of users are satisfy ,and seem to be prefer the online rating scale for the life distress inventory than the original paper-and-pencil version.

## 5.3 Weka's Result

Clustering and classification techniques were used in this research as follows :

Data Mining techniques can be used to cluster the data and then to create classification decisions from the clusters. *SimpleKMeans* was used as the main algorithm in this study because it is simple and its results are easy to interpret. In this study ,we evaluate the performance of the classifiers generated for different number of clusters (the *FP rates* were low; the *success rate*, the *TP rates* or the *precisions* were high).

The classification techniques used in this research are *J48. J48*, an implemented version of C4.5 in Weka, is one of the most extensively researched Data Mining methods. This algorithm constructs decision trees in a divide-and-conquer manner. The efficiency of existing decision tree algorithms, such as C4.5, has been established for relatively small datasets.

The data used in the experiments in this study include Gender, Status, Age, Education, Occupation and Result of testing.

In this research is divided into two experiments

- **Clustering Experiment 1**

The clustering experiments 1 can not be divided respondents whose status is married. The algorithm is applied to each dataset seven times, with the number of cluster being 2, 3, 4, 5, 6, 7 and 8. In this research,we evaluate the performance of the classifiers generated for different number of clusters (the *FP rates* were low; the *success rate*, the *TP rates* or the *precisions* were high). It was found that numCluster = 3 was chosen base on result are described in chapter 4.

- **Clustering Experiment 2**

The clustering experiments 2 can be divided respondents whose status is married. The algorithm is applied to each dataset four times, with the number of cluster being 9, 10, 11 and 12 (numClusters = 9, 10, 11 or 12) . In this research,we evaluate the performance of the classifiers generated for different number of clusters (the *FP rates* were low; the *success rate*, the *TP rates* or the *precisions* were high). It was found that numCluster = 9 was chosen. Summarized results are shows in Table 5.1.

**Table 5.1** The cluster's result of the Clustering experiment 2 is described as follow:

Number of iterations: 3

Within cluster sum of squared errors: 406.0

Missing values globally replaced with mean/mode

Cluster centroids:

| | | Cluster | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Attribute | Full Data | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| | (308) | (84) | (39) | (35) | (27) | (23) | (4) | (71) | (20) | (5) |
| Gender | Female | Female | Female | Male | Male | Female | Female | Female | Male | Male |
| Status | Single | Single | Single | Single | Single | Single | Single | Single | Single | Married |
| Age | 26-45 | 19-25 | 19-25 | 26-45 | 26-45 | 19-25 | 19-25 | 26-45 | 26-45 | 19-25 |
| Education | Bachelor | Bachelor | Bachelor | Bachelor | Bachelor | Bachelor | Diploma | Bachelor | Bachelor | Bachelor |
| Occupation | Government | Government | Students | Freelance | Government | Students | Government | Government | Employees | Employees |
| Output | Little stress | Little stress | Little stress | Little stress | Normal | Moderate stress | Little stress | Normal | Moderate stress | Severe stress |

| Clustered Instances | |
|---|---|
| 0 | 84 ( 27%) |
| 1 | 39 ( 13%) |
| 2 | 35 ( 11%) |
| 3 | 27 ( 9%) |
| 4 | 23 ( 7%) |
| 5 | 4 ( 1%) |
| 6 | 71 ( 23%) |
| 7 | 20 ( 6%) |
| 8 | 5 ( 2%) |

From Table 5.1, when the results showed that respondents in Cluster 1 and 4, which has characteristics similar but not identical effects of stress. Respondents in cluster1 have little stress but respondents in cluster 4 have moderate stress.

In this study, the user must register as a member before answer questionnaire about The Life Distress Inventory. The information in the registration will ask for personal information, which include Gender , Status ,Weight , Height , Age, Blood Group , Education , Occupation and Email address. The results from the Clustering experiment 2 which showed that although the characteristics of the respondents are the same, but the effect of stress is different.Because the people will react differently to stress even in the same situation [28] .Thus, in the personal information should be added questions about respondents' stress at that time.

Are you having stress now ?

◯ No

◯ Yes

If you have stress. What kind of stress is it ? (Choose at least one answer)

- Marriage
- Education and occupation
- Activity outside family
- Personal life and family
- Satisfaction in life

# CHAPTER VI
# CONCLUSION AND RECOMMENDATION

This chapter describes the conclusion and the recommendation for future works.

## 6.1 Conclusion

This research has an extension to the development of online rating scale for The Life Distress Inventory computerized version which originally was in form of paper-pencil based. The first objective of this study was to develop online rating scale for The Life Distress Inventory version with database for solve limitation for paper-pencil based, the report can be generated for researchers, specialists and clinician in psychology fields. The second objectives was using data mining techniques to study hidden knowledge's about relationship with personal information and The Life Distress Inventory. Process of this research was separated to development The Life Distress Inventory computerized version phase and analysis data phase by using data mining technique.

In first part, the development of online rating scale for The Life Distress Inventory computer based composed to design structure of application, to create web page and evaluate performance of online rating scale for The Life Distress Inventory version. In design phase, requirement from user interview was based to design structure of application which consists of Data Flow Diagram, ER Diagram and Data Dictionary. In creating phase was development web page by using Microsoft SQL as database,ASP.NET is server side programming. In evaluate phase, this application was evaluated by comparing the limitations of paper-pencil based with computer based. Besides, the satisfaction of respondents were evaluated by interview respondents. Totally, online rating scale for The Life Distress Inventory version is able to solve all limitations of paper-pencil based questionnaire which includes geographical limitation, time consuming and lack of database. Moreover, respondents's level of satisfaction in using online rating scale for The Life Distress Inventory is very satisfied range.

In the second part of this paper, in the analysis process, this study used Weka to analyze data. *SimpleKMeans* algorithm was used to cluster The Life Distress Inventory data. Results showed that the optimal number of segments can vary from *2* to 12 clusters, depending on the chosen information criterion. Thence, *J48* decision tree algorithm was used to classify the classes, with *cluster* being the target attribute.

In this research is divided into two experiments

- **Clustering Experiment 1**

The clustering experiments 1 can not be divided respondents whose status is married. The algorithm is applied to each dataset seven times, with the number of cluster being 2, 3, 4, 5, 6, 7 and 8. In this research,we evaluate the performance of the classifiers generated for different number of clusters (the *FP rates* were low; the *success rate*, the *TP rates* or the *precisions* were high). It was found that numCluster = 3 was chosen.

- **Clustering Experiment 2**

The clustering experiments 2 can be divided respondents whose status is married. The algorithm is applied to each dataset four times, with the number of cluster being 9, 10, 11 and 12 (numClusters = 9, 10, 11 or 12) . In this research,we evaluate the performance of the classifiers generated for different number of clusters (the *FP rates* were low; the *success rate*, the *TP rates* or the *precisions* were high). It was found that numCluster = 9 was chosen.

In the conclusion, Life Distress Inventory computer based can reduce limitations of paper-pencil based and the respondents had high satisfied to Life Distress Inventory computerized version. Moreover, knowledge discovery approach used clustering and decision tree to extract some hidden knowledge from the Life Distress Inventory databank. The extracted rules must be validated to ensure a satisfactory level of reliability. Then, the fusions of common and unique rules are performed to summarize the discovered knowledge. Once the knowledges are interpreted, the domain expert approval or further proof is required. Finally, the knowledge will be documented. Our research is currently proceeding with the domain expert approval.

## 6.2 Recommendation

6.2.1 Clustering and classification were only one data mining techniques. However knowledge in databank does not end, the other technique in data mining can find information. There should to use the other technique to analysis database such as association rules, for find the other hidden knowledges.

6.2.2 The development paper-pencil-based questionnaire to computer-based had more advantages. It can solve limitations of paper-pencil based such as geographical limitation, time consuming, cost Inefficiency and lacking of database. Therefore should to develop other computerized version of the other psychological tests, and find more innovative methods to produce the program and enhance the presentation.

# REFERENCES


1. สถาบันวิจัยประชากรและสังคม มหาวิยาลัยมหิดล. สุขภาพคนไทย 2552: เพื่อสุขภาวะแห่งมวลมนุษย์ หยุดความรุนแรง. พิมพ์ครั้งที่ 1. นครปฐม : อมรินทร์พรินติ้งแอนด์พับลิชชิ่ง จำกัด (มหาชน); 2552.
2. The Information and Public Relations office Ministry of Public Health [online]. Availablefrom:http://www.moph.go.th/show_hotnew.php?idHot_new=25942. [Accessed, 2010].
3. สุชีรา ภัทรายุตวรรตน์, เธียรชัย งามทิพย์วัฒนา, กนกรัตน์ สุขะตุงคะ, สมพร จรัสเจริญวิทยาและมาลัย เฉลิมชัยนุกูล. แบบสำรวจสถานการณ์ที่ก่อให้เกิดความเครียด. วารสารจิตวิทยาคลินิก. 2542; 30(2): 24-35.
4. ทรัพยากรด้านสุขภาพจิต [online] . Available from: http://www.dmh.moph.go.th/r1.asp. [Accessed, 2010].
5. สุชีรา ภัทรายุตวรรตน์. คู่มือการวัดทางจิตวิทยา. พิมพ์ครั้งที่3. กรุงเทพฯ : สำนักพิมพ์เมดิคัลมีเดีย;2548.
6. Brian Luke Seaward. Managing Stress: principle and strategies for health and well-being.5th ed.United States of America:Jones and Bartlett Publishers; 2006.
7. Lazarus, R.S. & Folkman S. (1984). Stress, appraisal and coping. New York : Springer Publishing company,Inc.
8. Wallace, J. Macdonald.(1978). Living With Stress. Nursing Time.74. 457-458.
9. ชูทิตย์ ปานปรีชา. (2534). จิตวิทยาทั่วไป. นนทบุรี : มหาวิทยาลัยสุโขทัยธรรมมาธิราช.
10. Janis I.L. Psychological stress. New York: John Wiley & sons,In; 1952.
11. กรมสุขภาพจิต กระทรวงสาธารณสุข.คู่มือการดำเนินงานในคลินิกคลายเครียด. นนทบุรี : กรมสุขภาพจิต กระทรวงสาธารณสุข; 2541.

12. Joel E. Dimsdale, Michael Irwin, Francis J. Keefe, Murray B. Stein. Stress and Psychiatry.In:Benjamin J Sadock, Virginia A. Sadock,editors.Kaplan & Sadock's Comprehensive Textbook of Psychiatry. NYU School of Medicine, New York.

13. Holmes T. H., and Rahe R.H. The Social Readjustment Rating Scale. Journal of Psychosomatic Research. 1967; 11 : 213-218.

14. Ronald R. Plew and Ryan Stephens. Sams Teach Yourself SQL in 24 Hours. United States of America: Sams Publishing;2002.

15. Cristian Darie and Wyatt Barnett. Build Your Own ASP.NET 3.5 Web Site Using C# & VB. United States of America: SitePoint Pty. Ltd;2008.

16. Han J. and Kamber M. Data Mining : Concepts and Techniques. Morgan Kaufmann. New York; 2001.

17. Dinesh P. Mehta and Sartaj Sahni. Handbook of data structures and applications. United States of America: A CRC Press Company;2004.

18. Groth, R. Data Mining : Competitive Advantage, Prentice-Hall. Upper Saddle River. NJ; 2000.

19. Groth R. Data Mining : A Hands-On Approach for Business Professionals. Upper Saddle River. NJ : Prentice Hall PTR; 1998.

20. Srisuriyachai, Sunida and Marukatat, Rangsipan. Finding Phenomenon Patterns using Decision Trees: A Case Study of Traffic Accidents in Nakorn Pathom. ECTI Transactions on Computer and Information Technology (ECTI-CIT). Special section on papers selected from NCSEC 2006 that will be published in May 2007.

21. Witten I. H. and Frank E. Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations. Morgan Kaufmann Publishers;2000.

22. Carol V. Brown, Heikki Topi. Information systems management handbook 8 th EDITION. United States of America : CRC Press LLC;2003.

23. สุนิดา ศรีสุริชัย และ รังสิพรรณ มฤคทัต. การหารูปแบบของปรากฏการณ์โดยใช้ต้นไม้การตัดสินใจ:กรณีศึกษาอุบัติเหตุการจราจรในจังหวัดนครปฐม.การประชุมวิชาการวิทยาการคอมพิวเตอร์และวิศวกรรมคอมพิวเตอร์แห่งชาติ ครั้งที่ 10. มหาวิทยาลัยขอนแก่น; 2549.

24. Wiyada Nintranon. The Development of The Life Distress Inventory Computerized version. [M.Sc. Thesis (Clinical Psychology)]. Bangkok:Faculty of Graduate Studies, Mahidol University; 2007.

25. Shannon L. Currie,Patrick J. McGrath,Victor Day. Development and usability of an online CBT program for symptoms of moderate depression, anxiety, and stress in post-secondary students. Computers in Human Behavior 2010; 26: 1419–1426.

26. Erik Hedman, et al. Internet administration of self-report measures commonly used in research on social anxiety disorder: A psychometric evaluation. Computers in Human Behavior 2010; 26: 736–740.

27. Weka: Data Mining Software in Java, University of Waikato, New Zealand [online]. Available form: http://www.cs.waikato.ac.nz. [Accessed, 2010].

28. ความเครียด..บางอย่างที่คุณควรรู้[online] **.** Available

from: http://www.dmh.moph.go.th/news/view.asp?id=912. [Accessed, 2010].

# APPENDICES

# APPENDIX A

ตัวอย่างแบบประเมิน "แบบวัดสถานการณ์ที่ก่อให้เกิดความเครียด"
( Life Distress Inventory)

พัฒนาโดย

รศ. ดร. สุชีรา ภัทรายุตวรรตน์ และคณะ
ภาควิชาจิตเวชศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล

**แบบวัดสถานการณ์ที่ก่อให้เกิดความเครียด**

**(Life Distress Inventory)**

**<u>ผู้พัฒนา</u>: สุชีรา ภัทรายุตวรรตน์ และคณะ**

ภาควิชาจิตเวชศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล

<u>Source</u>: วารสารสมาคมนักจิตวิทยาคลินิกไทย 2543; 45(3):237-50.

หากประสงค์จะนำแบบวัด ฉบับนี้ ไปใช้เพื่อการศึกษาวิจัย โปรดติดต่อที่

รศ.ดร. สุชีรา ภัทรายุตวรรตน์ ภาควิชาจิตเวชศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล

เขตบางกอกน้อย กรุงเทพฯ 10700

<u>ลักษณะเครื่องมือ</u>: มาตรประเมินแบบเรียงอันดับ (ordinal rating scale)

แบบ 5มาตร จำนวน 20 ข้อ

<u>จุดประสงค์</u> : แบบประเมินปัญหาที่ก่อให้เกิดความเครียดในประชากรทั่วไป

<u>ผู้รับการทดสอบ</u> : ทำการประเมินตนเอง

<u>การทดสอบ</u> : บุคคลทั่วไปที่สนใจ

ใช้เวลา 10 นาที

การให้คะแนน ให้คะแนนเป็นช่วง 0 – 4

(คะแนนรวมเทียบกับตารางที่ผู้วิจัยสร้างขึ้น โดยมีจุดตัดที่ 18 คะแนน)

<u>ตัวแปรที่วัด</u> : ภาวะความเครียด 5 ระดับคือ

1) ก่อให้เกิดความเครียดมาก

2) ก่อให้เกิดความเครียดในระดับปานกลาง

3) ก่อให้เกิดความเครียดเล็กน้อย

4) เกณฑ์ปกติ

5) ต่ำกว่าเกณฑ์ปกติอย่างมาก

<u>คุณภาพ</u> : Construct validity, discriminace validity,

Cronbach’s Alpha coefficient =0.84

Split-haft coefficient =0.88

<u>การนำไปใช้</u> : แบบคัดกรองปัญหาที่ก่อให้เกิดภาวะความเครียด

**แบบวัดสถานการณ์ที่ก่อให้เกิดความเครียดในชีวิตประจำวัน**

**คำชี้แจง** ข้อคำถามต่อไปนี้เป็นประสบการณ์ต่างๆในชีวิตประจำวันที่ท่านประสบจะให้ท่านทำการประเมินระดับความเครียดที่เกิดขึ้นกับการปรับตัวในแต่ละด้านซึ่งแบ่งออกเป็น 5 ระดับดังนี้

**4** **ก่อให้เกิดความเครียดมากที่สุด**

**3** **ก่อให้เกิดความเครียดมาก**

**2** **ก่อให้เกิดความเครียดในระดับปานกลาง**

**1** **ก่อให้เกิดความเครียดเพียงเล็กน้อย**

**0** **ไม่ก่อให้เกิดความเครียดเลย**

| **สภาพการณ์** | **มาก ที่สุด** | **มาก** | **ปานกลาง** | **เล็กน้อย** | **ไม่มี** |
|---|---|---|---|---|---|
| 1.ปัญหาการเจ็บป่วยของสมาชิกในครอบครัว | | | | | |
| 2.การปรับตัวทางเพศ | | | | | |
| 3.สัมพันธภาพกับคู่ครอง/คนรัก | | | | | |
| 4.สัมพันธภาพกับลูกๆ/พ่อแม่ | | | | | |
| 5.สัมพันธภาพกับญาติๆ | | | | | |
| 6.ชีวิตความเป็นอยู่โดยทั่วไปภายในบ้าน | | | | | |
| 7.สถานะทางการเงิน | | | | | |
| 8.การศึกษา | | | | | |
| 9.การว่างาน | | | | | |
| 10.บรรยากาศและสภาพแวดล้อมในที่ทำงาน | | | | | |
| 11.การใช้เวลาว่าง | | | | | |
| 12.การบริหารเวลา | | | | | |
| 13.สุขภาพทางด้านร่างกาย | | | | | |
| 14.สุขภาพทางด้านจิตใจ | | | | | |
| 15.ความเป็นอิสระในชีวิตส่วนตัว | | | | | |
| 16.ความพึงพอใจในชีวิต | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| 17.การคาดหวังเกี่ยวกับอนาคต | | | | | |
| 18.การปรับตัวกับเพื่อนๆ | | | | | |
| 19.การปรับตัวทางสังคม | | | | | |
| 20.ระดับความเครียดโดยรวมในขณะนี้ | | | | | |

**เกณฑ์ปกติ (Norms)**

| คะแนนที่ก่อให้เกิดความเครียดที่เกี่ยวข้องในชีวิตประจำวัน | ระดับความเครียด |
|---|---|
| 34 - 80 | ก่อให้เกิดความเครียดมาก |
| 26 - 33 | ก่อให้เกิดความเครียดในระดับปานกลาง |
| 19 - 25 | ก่อให้เกิดความเครียดเล็กน้อย |
| 4 - 18 | เกณฑ์ปกติ |
| 0 - 3 | ต่ำกว่าเกณฑ์ปกติอย่างมาก |

# APPENDIX B

# WEKA'S RESULTS

**The classification profile of each cluster (numCluster = 3) in Life Distress Inventory dataset**

```
=== Run information ===

Scheme:       weka.classifiers.trees.J48 -R -N 3 -Q 1 -M 2
Relation:     STRESS_LDI-weka.filters.unsupervised.attribute.Remove-R3-4,6-weka.filters.unsupervised.attribute.AddCluster-Wweka.clusterers.SimpleKMeans -N 3 -A "weka.core.EuclideanDistance -R first-last" -I 500 -S 10
Instances:    308
Attributes:   7
              Gender
              Status
              Age
              Education
              Occupation
              Output
              cluster
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

J48 pruned tree
------------------
Gender = Male
|   Age = 19-25: cluster2 (16.0/2.0)
|   Age = 26-45: cluster3 (40.0)
|   Age = 46-60: cluster3 (2.0)
Gender = Female
|   Age = 19-25
|   |   Occupation = Unemployed: cluster2 (1.0)
|   |   Occupation = Students: cluster2 (33.0)
|   |   Occupation = Employees: cluster2 (6.0)
|   |   Occupation = Government service: cluster1 (16.0)
|   |   Occupation = Freelance: cluster2 (2.0)
|   |   Occupation = Private business: cluster2 (0.0)
|   Age = 26-45: cluster1 (77.0)
|   Age = 46-60: cluster1 (13.0)

Number of Leaves  :   11
```

Size of the tree : 15

Time taken to build model: 0.08 seconds

=== Stratified cross-validation ===
=== Summary ===

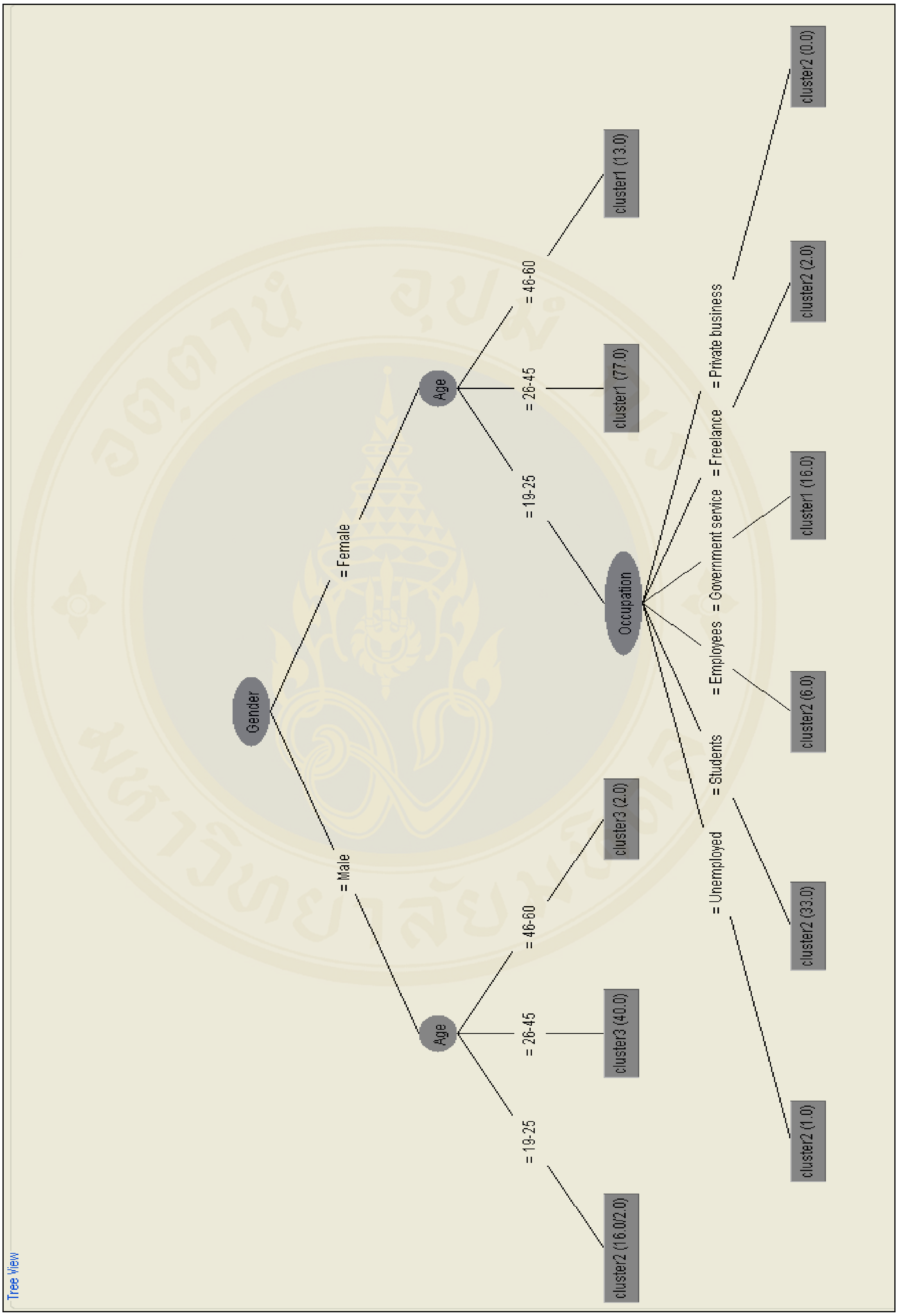
| | | |
|---|---|---|
| Correctly Classified Instances | 302 | 98.0519 % |
| Incorrectly Classified Instances | 6 | 1.9481 % |
| Kappa statistic | 0.9683 | |
| Mean absolute error | 0.0158 | |
| Root mean squared error | 0.1017 | |
| Relative absolute error | 3.8503 % | |
| Root relative squared error | 22.4646 % | |
| Total Number of Instances | 308 | |

=== Detailed Accuracy By Class ===

| | TP Rate | FP Rate | Precision | Recall | F-Measure | ROC Area | Class |
|---|---|---|---|---|---|---|---|
| | 0.987 | 0.013 | 0.987 | 0.987 | 0.987 | 0.997 | cluster1 |
| | 0.976 | 0.018 | 0.953 | 0.976 | 0.965 | 0.994 | cluster2 |
| | 0.969 | 0 | 1 | 0.969 | 0.984 | 0.991 | cluster3 |
| Weighted Avg. | 0.981 | 0.012 | 0.981 | 0.981 | 0.981 | 0.995 | |

=== Confusion Matrix ===

| a | b | c | <-- classified as |
|---|---|---|---|
| 157 | 2 | 0 | a = cluster1 |
| 2 | 82 | 0 | b = cluster2 |
| 0 | 2 | 63 | c = cluster3 |


Tree View
Gender
= Female
= Male
Age
= 19-25
= 26-45
= 46-60
cluster1 (13.0)
cluster1 (77.0)
Occupation
= Private business
= Freelance
= Government service
= Employees
= Students
= Unemployed
cluster2 (0.0)
cluster2 (2.0)
cluster1 (16.0)
cluster2 (6.0)
cluster2 (33.0)
cluster2 (1.0)
Age
= 19-25
= 26-45
= 46-60
cluster3 (2.0)
cluster3 (40.0)
cluster2 (16.0/2.0)

**The classification profile of each cluster (numCluster = 9) in Life Distress Inventory dataset**

```
=== Run information ===

Scheme:       weka.classifiers.trees.J48 -R -N 3 -Q 1 -M 2
Relation:     STRESS_LDI-weka.filters.unsupervised.attribute.Remove-R3-4,6-weka.filters.unsupervised.attribute.AddCluster-Wweka.clusterers.SimpleKMeans -N 9 -A "weka.core.EuclideanDistance -R first-last" -I 500 -S 10
Instances:    308
Attributes:   7
              Gender
              Status
              Age
              Education
              Occupation
              Output
              Cluster

Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

J48 pruned tree
------------------

Age = 19-25
|   Occupation = Unemployed: cluster1 (3.0/1.0)
|   Occupation = Students: cluster2 (37.0)
|   Occupation = Employees
|   |   Gender = Male: cluster7 (5.0)
|   |   Gender = Female: cluster1 (6.0/1.0)
|   Occupation = Government service: cluster1 (17.0/1.0)
|   Occupation = Freelance: cluster3 (5.0/1.0)
|   Occupation = Private business: cluster1 (1.0)
Age = 26-45
|   Gender = Male
|   |   Occupation = Unemployed: cluster3 (0.0)
|   |   Occupation = Students: cluster3 (1.0)
|   |   Occupation = Employees
|   |   |   Status = Single: cluster3 (10.0)
|   |   |   Status = Married: cluster7 (2.0)
|   |   |   Status = Widowed: cluster3 (1.0)
|   |   |   Status = Divorce: cluster3 (0.0)
|   |   Occupation = Government service: cluster4 (14.0/2.0)
|   |   Occupation = Freelance: cluster3 (4.0)
|   |   Occupation = Private business: cluster3 (5.0)
|   Gender = Female
```

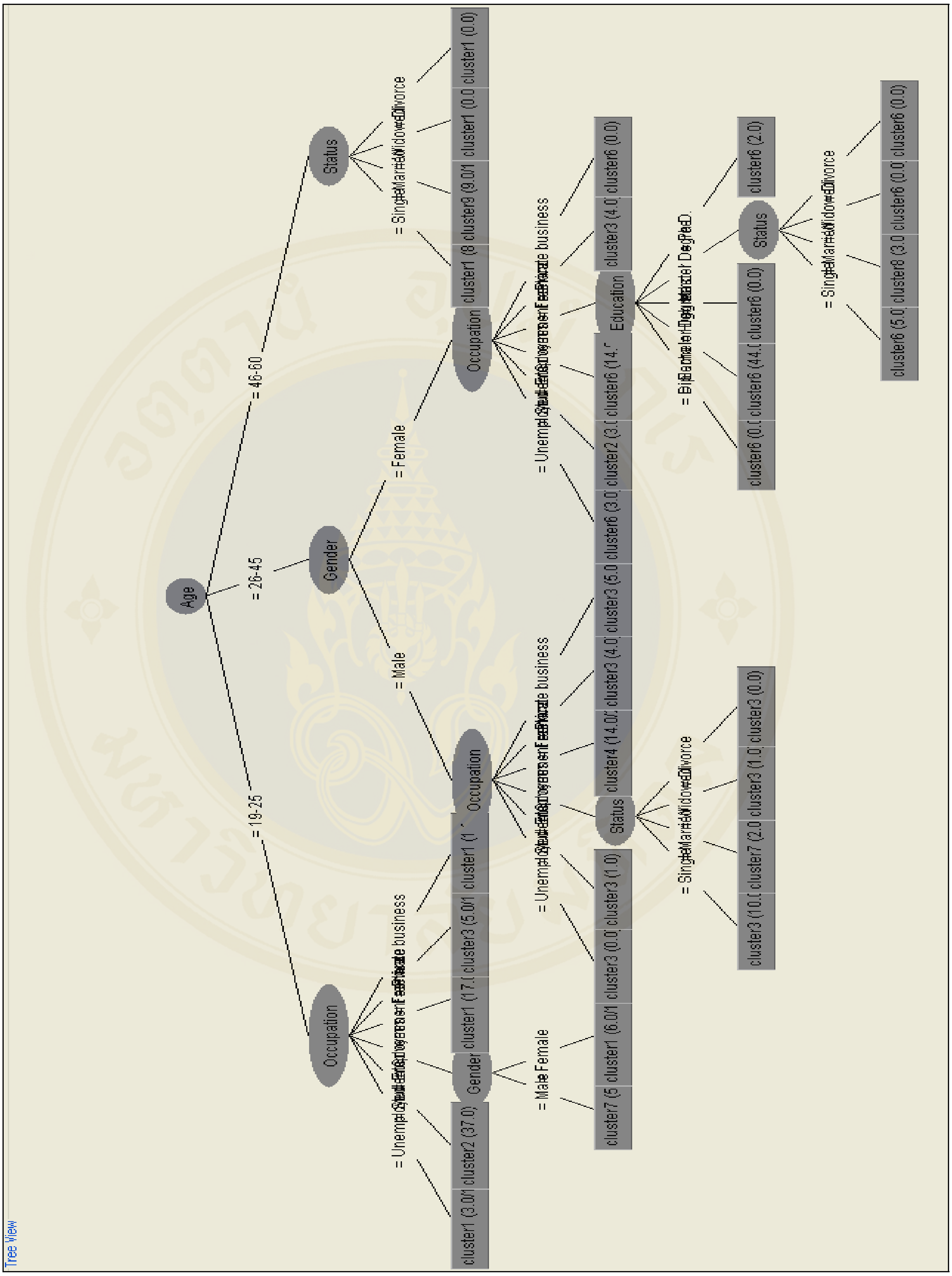
```
|  |  Occupation = Unemployed: cluster6 (3.0)
|  |  Occupation = Students: cluster2 (3.0)
|  |  Occupation = Employees: cluster6 (14.0)
|  |  Occupation = Government service
|  |  |  Education = Diploma: cluster6 (0.0)
|  |  |  Education =  Bachelor Degree: cluster6 (44.0)
|  |  |  Education = High school: cluster6 (0.0)
|  |  |  Education =  Master Degree
|  |  |  |  Status = Single: cluster6 (5.0)
|  |  |  |  Status = Married: cluster8 (3.0)
|  |  |  |  Status = Widowed: cluster6 (0.0)
|  |  |  |  Status = Divorce: cluster6 (0.0)
|  |  |  Education = Ph.D.: cluster6 (2.0)
|  |  Occupation = Freelance: cluster3 (4.0)
|  |  Occupation = Private business: cluster6 (0.0)
Age = 46-60
|  Status = Single: cluster1 (8.0)
|  Status = Married: cluster9 (9.0/1.0)
|  Status = Widowed: cluster1 (0.0)
|  Status = Divorce: cluster1 (0.0)

Number of Leaves  : 33

Size of the tree :       43


Time taken to build model: 0.02 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        293               95.1299 %
Incorrectly Classified Instances       15                4.8701 %
Kappa statistic                        0.9391
Mean absolute error                    0.0177
Root mean squared error                0.1065
Relative absolute error                9.8852 %
Root relative squared error           35.6786 %
Total Number of Instances             308
```

=== Detailed Accuracy By Class ===

| | TP Rate | FP Rate | Precision | Recall | F-Measure | ROC Area | Class |
|---|---|---|---|---|---|---|---|
| | 0.92 | 0.012 | 0.939 | 0.92 | 0.929 | 0.981 | cluster1 |
| | 1 | 0.004 | 0.983 | 1 | 0.992 | 1 | cluster2 |
| | 1 | 0.019 | 0.898 | 1 | 0.946 | 0.992 | cluster3 |
| | 1 | 0.01 | 0.864 | 1 | 0.927 | 0.992 | cluster4 |
| | 0 | 0 | 0 | 0 | 0 | 0.479 | cluster5 |
| | 1 | 0 | 1 | 1 | 1 | 1 | cluster6 |
| | 0.667 | 0 | 1 | 0.667 | 0.8 | 0.827 | cluster7 |
| | 0.556 | 0.003 | 0.833 | 0.556 | 0.667 | 0.797 | cluster8 |
| | 1 | 0.007 | 0.846 | 1 | 0.917 | 0.994 | cluster9 |
| Weighted Avg. | 0.951 | 0.006 | 0.944 | 0.951 | 0.944 | 0.977 | |

=== Confusion Matrix ===

| a | b | c | d | e | f | g | h | i | <-- classified as |
|---|---|---|---|---|---|---|---|---|---|
| 46 | 1 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | a = cluster1 |
| 0 | 59 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | b = cluster2 |
| 0 | 0 | 44 | 0 | 0 | 0 | 0 | 0 | 0 | c = cluster3 |
| 0 | 0 | 0 | 19 | 0 | 0 | 0 | 0 | 0 | d = cluster4 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | e = cluster5 |
| 0 | 0 | 0 | 0 | 0 | 101 | 0 | 0 | 0 | f = cluster6 |
| 1 | 0 | 3 | 0 | 0 | 0 | 8 | 0 | 0 | g = cluster7 |
| 0 | 0 | 0 | 3 | 0 | 0 | 0 | 5 | 1 | h = cluster8 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 11 | i = cluster9 |


Tree View
Age
= 19-25
= 26-45
= 46-60
Occupation
Gender
Status
= Male
= Female
Education
cluster2 (37.0)
cluster6 (2.0)
cluster6 (44.1
cluster6 (0.0)
cluster1 (0.0)

# BIOGRAPHY

| | |
|---|---|
| **NAME** | Miss. Suwannee Suksawas |
| **DATE OF BIRTH** | May 9 1982 |
| **PLACE OF BIRTH** | Ranong, Thailand |
| **INSTITUTIONS ATTENDED** | Mahidol University, 2001- 2005: Bachelor of Nursing |
| | Mahidol University, 2007 - 2011: Master of Science (Technology of Information System Management) |
| **POSITION & OFFICE** | 2006 – Present Central Chest Disease Institute of Thailand, Nonthaburi. |
| | Position: Nurse |
| | Tel : 02-5803423 |
| | E-mail : Suwannee.cpl@gmail.com |
| | Botanical.mu@hotmail.com |
| **HOME ADDRESS** | 121/1Moo 3 Hadsompan,Muang,Ranong 85000 |
| | Tel : 077-862451 |