School of Biomedical Engineering

University of New South Wales

Machine Learning Analysis of Multiparameter Home Telecare Clinical Measurement Data for Patient Health Stratification

A Project presented for the degree

Master of Engineering in Biomedical Engineering

July 2008

Mohd Hanafi Mat Som

3205066

Coordinator: Prof. Nigel Lovell

Supervisor 1: Dr. Jim Basilakis

Supervisor 2: Dr. Stephen Redmond

ABSTRACT

A medical decision support system (DSS) is designed to assist clinician in monitoring patient's health by the means of providing reminders, advice as well as interpretation. This system is good enough in the sense of monitoring patient's health and improves the early prevention by the means of providing 'just-in-time' notifications for the best action to be taken. In Biomedical System Laboratory (BSL) of UNSW, there is a system as such described that provides information on the patient's conditions based on the risk they may have by analyzing the data entered in the database. But there is a need to improve current system to increase their performance. Score generated by the DSS is compared with the journal on patient's conditions entered manually by medical personnel. The objective is to see the reliability of the score generated by the DSS. Unfortunately, due to some problems on the data, the outcome of this study can't be used for reference. However, the method used can be repeated but with a better database storing the patient's information.

i

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to all people contributing to the completion of this project.

To all my supervisors, Prof. Nigel Lovell, Dr. Jim Basilakis and Dr. Stephen Redmond, for giving me this opportunity to involve in this project. Thank you for continuous encouragement, exceptional intelligence and above all giving me guidance during all this time. Thank you.

To the technical staff in Biomedical System Laboratory, Neil McKenzie for his assistance.

To my fellow students, especially those in the laboratory, for giving me support and ideas.

To my parents, for being a good listener to my problems and thank you for always being there for me.

Finally, to my sponsor, Ministry of Higher Education Malaysia and University of Malaysia Perlis, for giving me a chance to further my study in Australia.

Thank you.

LIST OF CONTENTS

iv

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

A medical decision support system (DSS) can be described as computer program that provides reminders, advice or interpretation specific to a given patient at a particular time [1]. They are also designed to assist clinicians in making decisions about patient care by providing assessments or prompts specific to individuals based on the characteristics provided in the knowledge based. This system is good for both consumer and clinicians by the means of providing 'just-in-time' notifications for the best action to be taken and access to relevant, evident-based information at the point of care respectively [2]. There are many other definitions of DSS and the scope it covers are varies but for the purpose of this report only restricted to the notification of patients conditions based by the DSS based on rules defined.

As for now in Biomedical System Laboratory (BSL) of UNSW, there is a system of DSS that provides information on the patient's conditions based on the risk they may have by the means of analysing the data provided in the database.

1.1 Objectives

This report is written to present the outcome of the study by comparing the output produce by the DSS with the data of 31 patients taken from a database that contains relevant information for each patient daily. The performances of the data were analysed using two models which are Linear Least Square Model and Linear Discriminant Classifier Model. Both methods will be further discussed on later chapter.

Chapter 2 basically will discuss briefly on steps and method used in this study. It includes the software and database used, the development process, features selected for the modelling purposes and also the measure of performance of the system. Chapter 3 will explain on the outcome of this study. Meanwhile chapter 4 will discuss on the result and outline some of the problems occur with recommendation on how to improve the system for future development.

Chapter 2

METHODOLOGY

2.1 Database and Software

Data from a total of 31 patients were used in this study. Patient data was derived from the Austin Hospital Remote Patient Monitoring Project, conducted by the HARP (Hospital Admissions Reduction Program) management team from the Austin Hospital in Melbourne. The duration of interest period is from 15th September 2006 until 4th March 2008 which total of 536 days. Due to time constrain, this study only implements the data from table that contains trends of the patient's score and some features. The details of the features will be discussed later on the next section.

SQL Server Management Express and Matlab have been used in this study. SQL Server Management Express is used mainly to provide GUI of the data in the database. It is also very helpful in creating and executing query to obtain data needed for the purpose of this study. Matlab on the other hand is the core software which does most of the job from translating the data into matrices until comparing the result of DSS with the data studied.

2.2 Development Process

The reason of this study is to see whether the journal entries supplied in the database are worth to use to conduct a bigger project to improve the DSS currently in use. The first phase of the study was to narrow down the journal entries into possible outcome we might be interested

3

since the journal contains too much information that is not related to the study. Three categories were suggested which are admission (class 1), visit (class 2) and status (class 3).

Patient who are admitted or already admitted into the hospital will be put into class 1 group. Meanwhile for every non-routine visit by general practitioner (GP) or to GP or by nurse will be classified as class 2. Whereas class 3 are for those patient's who are having worsening symptoms or start a medicine or any changes in medicine prescription.

The next step is to extract a few tables from database into matrices in Matlab program and then did some analysis graphically whether the journal entries data has some significant value with DSS result. This is quite an important step as it gives the whole idea which would verify the reliability of the data.

Three matrices had been produced from a few tables in the database and they are all plotted into the same graph as shown in figures below. In figure 2.1 and figure 2.2 the top (A), center (B) and bottom (C) of the plot are the data for classes, score from DSS and type of measurement done respectively. Whenever there is a value in A, the score given in B is observed to see the relationship. Type of measurement can be seen in C and should be reasonable. Say we have class 1 (admission) in A, the score shown in B should be more negative than usual. Otherwise the data might be unreliable and can't be used for further analysis. Having this plot analyzed and approved by Dr. Jim, this study was continued to the next step.

Using a table containing trends of the data and some features, this table then was translated into a matrix in Matlab. A matrix storing the 'Status Risk' was then used together with the matrix generated from trends table. 'Status Risk' is simply the class where the patient's belong to.
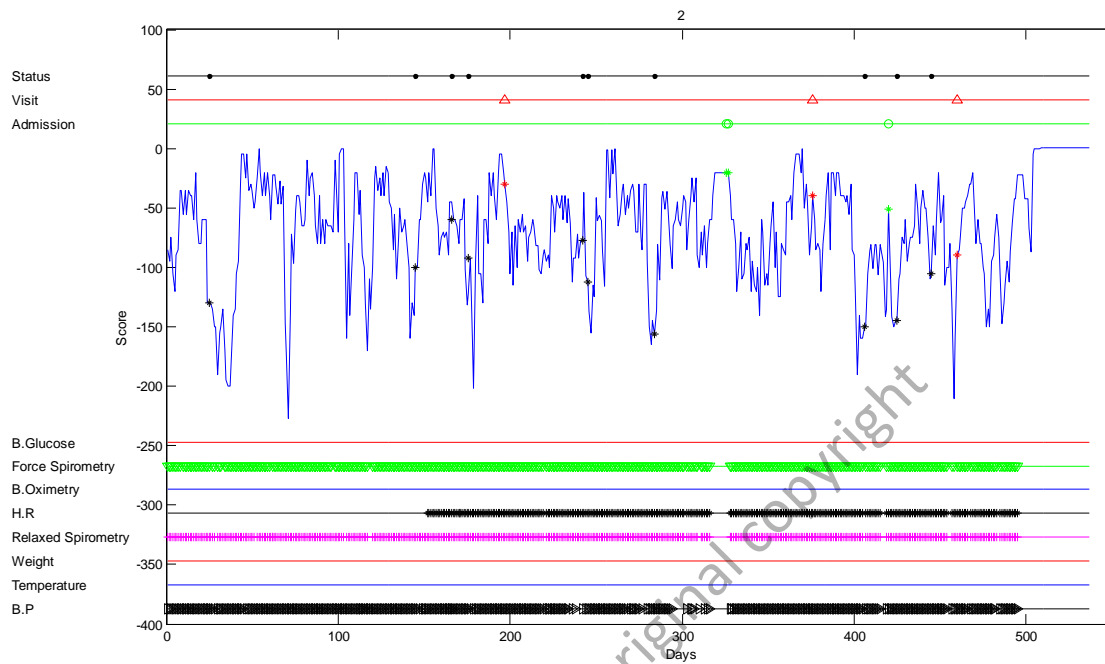
Figure 2.1: Example of visual observation for the reliability of the data for a particular patient.
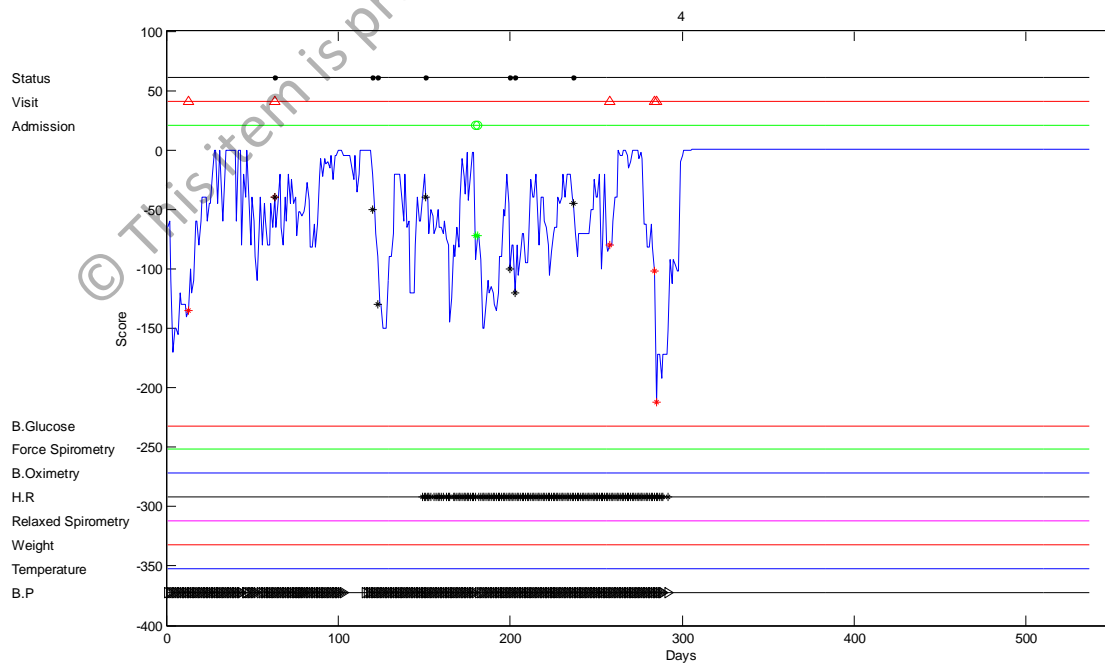
Figure 2.2: Another example of visual observation for the reliability of the data for a particular patient.

5

2.3 Features

Feature List	Slope	P value	R Square
blood sugar level	1	2	3
diastolic blood pressure	4	5	6
forced expiratory volume in 1 second	7	8	9
forced expiratory volume in 6 seconds	10	11	12
forced vital capacity	13	14	15
heart rate for the blood pressure	16	17	18
heart rate for the ecg	19	20	21
heart rate for the pulse oximtery	22	23	24
inspiratory capacity	25	26	27
respiratory rate	28	29	30
pulse oximetry	31	32	33
systolic blood pressure	34	35	36
temperature	37	38	39
Ti/Ttot (ratio of inspiratory time (Ti) over total breathing cycle duration (Ttot)) ie: fraction of the breathing cycle spent in inspiration	40	41	42
minute ventilation	43	44	45
tidal volume	46	47	48
weight	49	50	51

Table 2.1: Features used in the study which taken from the trends table in the database. Each feature

containing the data for slope, R squared and the P value.

Table 2.1 above shows the features used in this study. Every feature has their own P value, R square and slope taken from 2 weeks block of data. This study will only consider these parameters to determine the risk level and compare with the DSS result.

2.4   Models

In this study, 2 models were used to analyse the performance of the data. They are Linear Least Square Model (LLSM) and Linear Discriminant Classifier Model (LDCM).

2.4.1   Linear Least Square Overview

Linear least squares were used in this study primarily to fit a linear mathematical model to our data. The goals are to extract predictions from the sets of data and at the same time reducing the effect of measurement errors. In this method, the unknown parameters are estimated by finding numerical values for the parameters that minimize the sum of the squared deviations between the observed responses and the functional portion of the model. In other words, it is a problem of finding the estimate solution to an overdetermined system of linear equation.

For this study, estimation has been made using this method. This model requires a matrix, X, which have N rows and M subjects. Each row contains a number of features for each subject and a vector, r, containing the N 'gold-standard' risk status (class) scores. A set of weights, w, were determined to reduce the squared error between an estimate scores of the status risk, $\hat{r}$.

$$Xw = \hat{r},$$

and the 'true' status risk, r. Hence, the solution for this problem can be write as

$$w = X^+ r,$$

where X^+ is the pseudo-inverse of X.

7

2.4.2 Linear Discriminant Classifier Overview

Linear discriminant analysis (LDA) is one of the methods used in statistics and machine learning. It is used to find the best linear combination of selected variables that separate two or more classes of data by modelling them according to the difference between them [3]. This method also uses some searching to achieve the separation of classes. In our case, we have 3 classes say class 1, class 2 and class 3. The unique discriminant function can be obtained after doing training on the data.

2.5 Performance Measures

2.5.1 Root-Mean-Square Estimation Error

2.5.1.1 Cross-fold validation (leave one out)

The LLSM was trained using data from 31 subjects. Each patient has a number of data of their own within the period of interest. To avoid bias in the estimation, data from the ith subject is leave out of the training list during the model being trained. Once the training completed, the calculated weight, w_i, are used to calculate the approximate status risk for i^{th} subject using simple multiplication as follow

$$\hat{r} = x_i w_i,$$

where x_i denotes the features for each ith subject.

This training was repeated for all 31 subjects. Root-mean-squared of the error, *rmse* is calculated as a measure of performance.

$$rmse = \sqrt{\frac{\sum_{i=1}^{31}(\hat{r}_i - r_i)^2}{31}}$$

2.5.1.2 Feature selection

To minimise the root-mean-squared error, *rmse*, a search algorithm was employed to find for the optimal subset of features. This search algorithm was developed using a Sequential Forward Selection (SFS) which is a greedy algorithm that will test the performance of the system using only one feature at a time [4]. Suppose this selection will reduce the *rmse*, the algorithm will remember this selection by adding this feature to the selected set of features. This searching algorithm will continue to find the best feature among unselected features, then append to the first selected features and assessed. The combination with the best performance will be added to the set of selected features and the process will continue until no other features can be added to the selected set of features.

Then, when the selection cycle is done, the removal of the selected features is also considered using Sequential Backward Selection (SBS) which is similar with SFS except it removes the features from provided a set of features. This algorithm will continue until the best performance is achieved.

2.5.2 Cohen's Kappa Coefficient, κ

Cohen's Kappa Coefficient (κ) is a statistical measure of inter-rater agreement. In our case, the two raters are the score result from DSS and the outcome of risk status called 'Status Risk' from journal entries.

9

A search algorithm was used to find the optimal subset of features that will maximize the κ value. The algorithm work in the same manner as previously discussed above as they are using SFS and SBS.

Chapter 3

RESULT

This study implements two type of model for analyzing the performance of the data provided in journal entries as described earlier which are LLSM and LDCM. Both methods will have Specific Entries and General Entries result which implement only specific data available in journal entries and all the data available in DSS respectively. In our case, 314 and 7640 data for Specific Entries and General Entries result respectively.

3.1 Linear Least Square Model

3.1.1 Specific Entries Results

Figure 3.1 and 3.2 below shows the plot of the model using LLSM and the data used for this model is resulted from cross join the journal entries which have total of 31 patients with corresponding result of 31 patients from DSS. This gives 314 out of 7460 rows of data. The difference between both model are the values of the missing entries where model in figure 3.1 use zero in place of missing values meanwhile model in figure 3.2 replacing the missing entries by the mean value of each features respectively. Say we have 51 features; every missing entry in the feature 1 will be replaced by the mean of feature 1 and so on. This model classifies the data into 3 groups (1, 2 and 3).
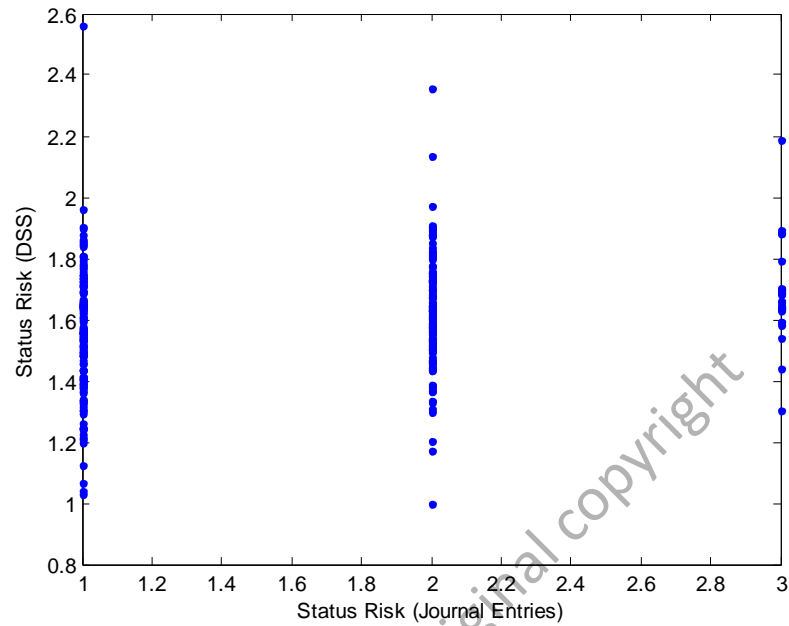
11

Figure 3.1: Specific Entries model using LLSM. Features that have no or missing values were

replace by zero.

Figure 3.2: Specific Entries model using LLSM. Features that have no or missing values were replace by

the mean of each features.

12

Model (Figure)	Features Selected (in order of selection)
3.1 (with zero value in place of missing entries)	49, 40, 23, 14, 8, 9, 37, 43, 4
3.2 (with mean value of respective features in place of missing value)	49, 40, 14, 23, 8, 9, 37, 43, 4

Table 3.1: Features selected for Specific Entries results

Table 3.1 above list out the features selected from total of 51 features for the Specific

Entries case. This selection is made by the features selection algorithm.

3.1.2 General Entries Results

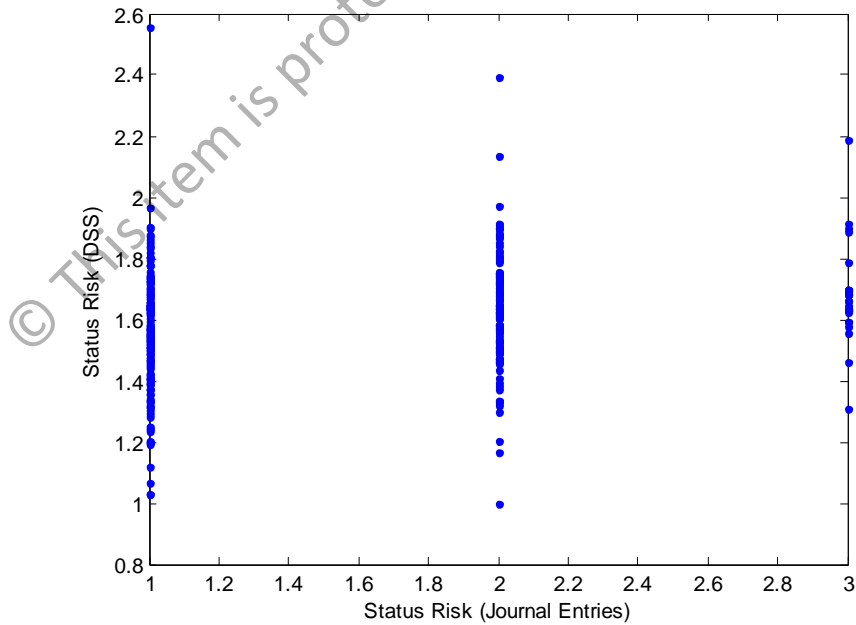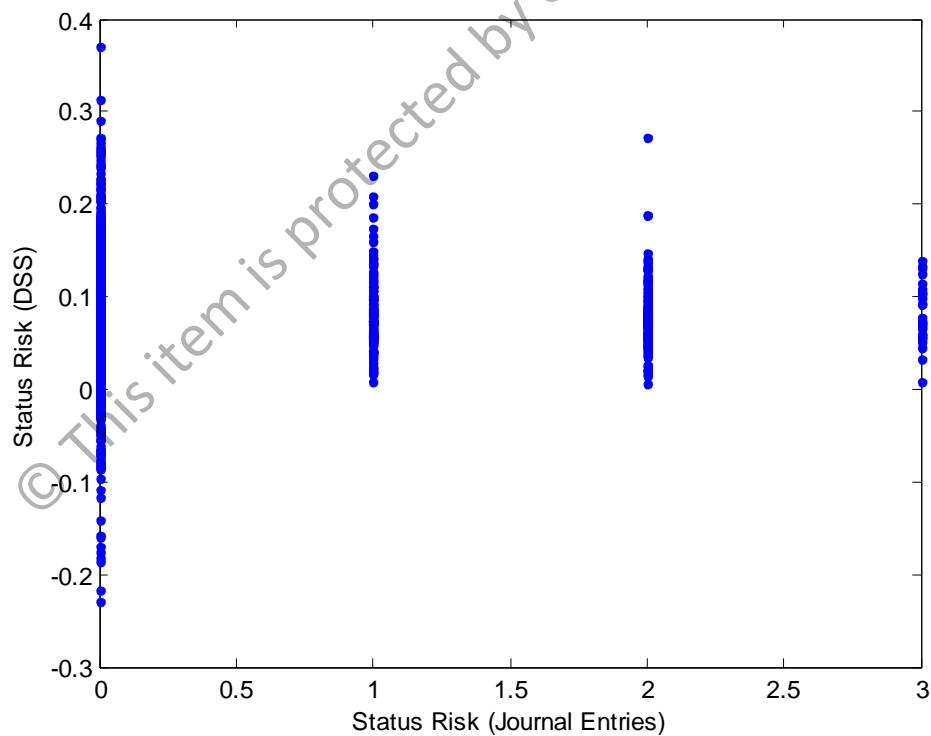*Figure 3.3: General Entries model using LLSM. Features that have no or missing values were replace by*

zero.
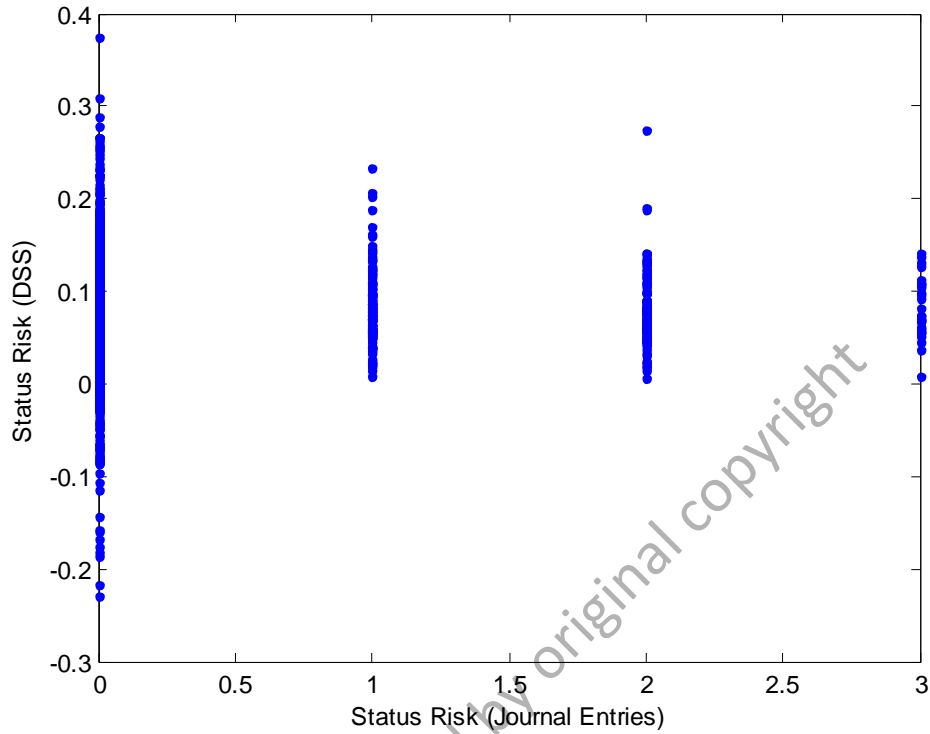
Figure 3.4: General Entries model using LLSM. Features that have no or missing values were replace by the mean of each features.

Figure 3.3 and 3.4 above shows the plot of the model using LLSM and the data used for this model is taken from all the entries of the patients which in this case 7460 rows of data. This model classifies the data into 4 groups (0, 1, 2 and 3).

Model (Figure)	Features Selected (in order of selection)
3.3 (with zero value in place of missing entries)	9, 13, 31, 25, 6, 19, 30, 22, 51, 16
3.4 (with mean value of respective features in place of missing value)	9, 13, 31, 25, 6, 51, 19, 30, 22, 16

Table 3.2: Features selected for General Entries results

3.2 Linear Discriminant Classifier Model

Tables below show the confusion tables for rater Status Risk (DSS) and rater Status Risk (Journal Entries). They will be referring to as rater A and rater B respectively from now on. The number in parenthesis is the fraction of overall total whereas the rest are the number of times rater A classified an observation as class *i* when rater B classified as class *j*.

3.2.1 Specific Entries Results

Status Risk (DSS)	Status Risk (Journal Entries)			
	1	**2**	**3**	**Total**
1	115 (0.366)	85 (0.271)	15 (0.048)	215 (0.685)
2	30 (0.096)	32 (0.102)	8 (0.025)	70 (0.223)
3	11 (0.035)	12 (0.038)	6 (0.019)	29 (0.092)
Total	156 (0.497)	129 (0.411)	29 (0.092)	314 (1)

Table 3.3: Confusion table for Specific Entries results. Missing entries in the journal was replaced by zero value and total numbers of data used are 314 for all 31 patients.

Status Risk (DSS)	Status Risk (Journal Entries)			
	1	**2**	**3**	**Total**
1	120 (0.382)	94 (0.299)	19 (0.061)	233 (0.742)
2	22 (0.070)	22 (0.070)	2 (0.006)	46 (0.146)
3	14 (0.045)	13 (0.041)	8 (0.025)	35 (0.111)
Total	156 (0.497)	129 (0.411)	29 (0.092)	314 (1)

Table 3.4: Confusion table for Specific Entries results. Missing entries in the journal was replaced by mean value with respect to each feature.

15

Confusion Table	Features Selected (in order of selection)
3.3 (with zero value in place of missing entries)	20, 27, 16
3.4 (with mean value of respective features in place of missing value)	20, 48

Table 3.5: Features selected for Specific Entries results of the confusion tables

3.2.2 General Entries Results

	Status Risk (Journal Entries)				
Status Risk (DSS)	**0**	**1**	**2**	**3**	**Total**
0	6784 (0.909)	143 (0.019)	120 (0.016)	27 (0.004)	7074 (0.948)
1	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
2	73 (0.010)	0 (0)	1 (0)	0 (0)	74 (0.10)
3	291 (0.039)	13 (0.002)	7 (0.001)	1 (0)	312 (0.042)
Total	7148 (0.958)	156 (0.021)	128 (0.017)	28 (0.004)	7460 (1)

Table 3.6: Confusion table for General Entries results. Missing entries in the journal was replaced by zero value and total numbers of data used are 7460 for all 31 patients.

	Status Risk (Journal Entries)				
Status Risk (DSS)	**0**	**1**	**2**	**3**	**Total**
0	6786 (0.910)	143 (0.019)	120 (0.016)	27 (0.004)	7074 (0.949)
1	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
2	71 (0.010)	0 (0)	1 (0)	0 (0)	74 (0.10)
3	291 (0.039)	13 (0.002)	7 (0.001)	1 (0)	312 (0.042)
Total	7148 (0.958)	156 (0.021)	128 (0.017)	28 (0.004)	7460 (1)

Table 3.7: Confusion table for General Entries results. Missing entries in the journal was replaced by mean value with respect to each feature.

Confusion Table	Features Selected (in order of selection)
3.6 (with zero value in place of missing entries)	1, 3
3.7 (with mean value of respective features in place of missing value)	1, 3

Table 3.8: Features selected for General Entries results of the confusion tables

3.2.3 Comparative κ value from the outcome

From the confusion tables above, we can find the Cohen's Kappa Coefficient, κ for each of them to measure the performances by using formula A.1, A.2 and A3 as shown in appendix A. The coefficient values are illustrated in table below.

Data Subject (Table)	κ value
3.3 (case A)	0.084
3.4 (case B)	0.069
3.6 (case A)	0.008
3.7 (case B)	0.008

Table 3.9: Table showing Cohen's Kappa Coefficient, κ for outcome of the study

17