

**IMPROVED HEALTH INFORMATION CONCERN THROUGH
DATA MINING OF MEDICAL RECORDS IN CAVITE
STATE UNIVERSITY INFIRMARY**

THESIS

ARVILYN ROGEL

VANESSA FUNNY M. PARAS

**College of Engineering and Information Technology
CAVITE STATE UNIVERSITY
Indang, Cavite**

March 2013

IMPROVED HEALTH INFORMATION CONCERN THROUGH DATA MINING OF MEDICAL RECORDS IN CAVITE STATE UNIVERSITY INFIRMARY

**Undergraduate Thesis
Submitted to the Faculty of the
College of Engineering and Information Technology
Cavite State University
Indang, Cavite**

**In partial fulfillment
of the requirements for the degree
Bachelor of Science in Computer Science**

**ARVILYN ROGEL
VANESSA FUNNY M. PARAS**



*Improved health information concern
through data mining of medical records in
005 7 R63 2013
1.5097*

March 2013

ABSTRACT

ARVILYN ROGEL and VANESSA FUNNY PARAS, Improved Health Information Concern through Data Mining of Medical Records in Cavite State University Infirmary , Undergraduate Thesis, Bachelor of Science in Computer Science, Cavite State University, Indang, Cavite, March 2013. Adviser: Mark Steve V. Poniente.

The study entitled “Improved Health Information Concern through Data Mining of Medical Records in Cavite State University Infirmary” was conducted to design and develop a Decision Support System for the Cavite State University Infirmary for easy managing of medical health records that can help the doctor in decision making of diagnoses to the patients through data mining that will produce a reasonable accuracy of predicting potential or possible diseases.

The proposed study incorporated the Apriori Algorithm which was designed to utilize the text mining method for clustering useful patterns form large volume of data to be analyzed. The software provides functionality that can predict potential and possible diseases based on useful patterns of data.

The proponents used Fourth Generation Technique in developing the software. These include four (4) phases; Requirements Gathering, Design Strategy, Implementation using Fourth Generation Language (4GL), and Testing. Adobe Photoshop CS3 was used in the design and Microsoft Visual Studio 2010- C# for prototyping.

The software was evaluated by patients/enrolled college students of Cavite State University-main campus and the Director, the doctor, nurses and other staff in the Infirmary on the following criteria: functionality, reliability, usability, efficiency, maintainability, and portability. The system passed the criteria and met all the objectives and requirements.

TABLE OF CONTENTS

	Page
BIOGRAPHICAL DATA	i
ACKNOWLEDGEMENT	ii
ABSTRACT.....	iv
LIST OF TABLES	v
LIST OF FIGURES.....	vii
LIST OF APPENDICES.....	viii
INTRODUCTION.....	1
Statement of the Problem	3
Theoretical Framework	5
Significance of the Study	8
Objectives of the Study	8
Time and Place of the Study.....	9
Scope and Limitation of the study	9
Definition of Terms	12
REVIEW OF RELATED LITERATURE.....	14
Related Literature.....	14

Related Studies	22
MATERIALS AND METHODS	26
Materials	26
Methodology.....	27
Breakdown of Respondents.....	29
Population and Sampling	31
Research Instrument.....	31
Data Collection.....	31
Statistical Analysis.....	31
RESULTS AND DISCUSSIONS.....	33
Problem Identification and Requirements Gathering.....	33
Design and Concept Specification.....	34
Screen Layout of the System.....	34
Implementation Using Fourth Generation Technique	39
Software Testing.....	42
Evaluation of the Software.....	43
Results of Evaluation	44
SUMMARY, CONCLUSIONS AND RECOMMENDATIONS.....	50
Summary	50

Conclusions	51
Recommendations.....	53
REFERENCES.....	54
APPENDICES.....	55

LIST OF TABLES

Table		Page
1	Breakdown of respondents	30
2	Mean score for functionality of the software	44
3	Mean score for reliability of the software	45
4	Mean score for usability of the software	46
5	Mean score for efficiency of the software	47
6	Mean score for maintainability of the software	48
7	Mean score for portability of the software	49

LIST OF FIGURES

Figures	Page
1 Theoretical framework of the proposed study	7
2 Knowledge Discovery in Database(KDD)	15
3 Fourth Generation Technique (4GT) methodology	25
4 Screen layout of log in form	34
5 Screen layout of welcome screen and main menu	35
6 Screen layout of view records menu	35
7 Screen layout of detailed report menu	36
8 Screen layout of add illness record menu	36
9 Screen layout of search student menu.....	37
10 Screen layout of graphical representation menu	37
11 Screen layout of medicine inventory menu	38
12 Screen layout of medicine supply menu	38
13 Screen layout of add new user account menu	39

LIST OF APPENDICES

Appendix 1	Page
Figures	
1	Data Flow Diagram (DFD) of the proposed system 56
2	Decision tree of the existing system..... 57
3	Use case diagram of the proposed system 58
Appendix 2	
Tables	
1	Frequency Distribution of the respondents perception for Functionality 63
2	Frequency Distribution of the respondents perception for Reliability 63
3	Frequency Distribution of the respondents perception for Usability 63
4	Frequency Distribution of the respondents perception for Efficiency 64
5	Frequency Distribution of the respondents perception for Maintainability 64
6	Frequency Distribution of the respondents perception for Portability 64

Appendix 3

Gantt Chart

1 Gantt Chart for the time schedule of the study 66

Appendix 4

Sample Evaluation Sheet

1 Sample Questionnaire..... 68

Appendix 5

Interview

1 Interview Report..... 71

Appendix 6

System Testing

1 Testing 73

Appendix 7

List of codes

1 Codes 74

IMPROVED HEALTH INFORMATION CONCERN THROUGH DATA MINING OF MEDICAL RECORDS IN CAVITE STATE UNIVERSITY INFIRMARY

Arvilyn Rogel

Vanessa Funny Paras

An undergraduate thesis manuscript submitted to the faculty of the Department of Information Technology, College of Engineering and Information Technology, Cavite State University, Indang, Cavite in partial fulfillment of the requirements for the degree Bachelor of Science in Computer Science with Contribution No. _____. Prepared under the supervision of Mr. Mark Steve V. Poniente.

INTRODUCTION

The healthcare industry is among the most information intensive industries. Medical information, knowledge and data keep growing on a daily basis. The Infirmary is considered as the only healthcare facility at Cavite State University. It handles massive amount of data and information on medical records of the people in the university. The Infirmary keeps all the health care information and its doctor perform several diagnoses that are important for students who want to be qualified for enrollment in their respective courses. And managing such large data and information is not an easy job and it is likewise a crucial task.

Data mining is a relatively new field of research whose major objective is to acquire knowledge from large amounts of data. Data mining, also called knowledge discovery in databases, in computer science, is the process of discovering interesting and useful patterns and relationships in large volumes of data. The field combines tools from statistics and artificial intelligence (such as neural networks and machine learning) with database management to analyze large digital collections, known as data sets. Data