

UNIVERSITY OF NAIROBI

SCHOOL OF COMPUTING AND INFORMATICS

**DECISION SUPPORT SYSTEM ON BAD DEBT
RECOVERY IN THE DEPOSIT PROTECTION
FUND BOARD - KENYA**

By

Simon Nyahe Waithaka

P56/7424/2005

August 2011

**Project submitted in partial fulfillment of the degree of Masters of
Science in Information Systems**

DECLARATION

THIS PROJECT IS MY ORIGINAL WORK AND HAS NOT BEEN SUBMITTED FOR A DEGREE IN
ANY OTHER UNIVERSITY

Signed:

Simon N. Waithaka

P56/P/7424/2005

Date:

**THIS PROJECT HAS BEEN SUBMITTED FOR EXAMINATION WITH MY APPROVAL AS A
UNIVERSITY SUPERVISOR**

Signed:

Muchemi L.

Lecturer

School of Computing and Informatics

University of Nairobi

Date:

ABSTRACT

The Deposit Protection Fund Board (DPFB), in particular Institutions in Liquidations division, is often plagued by a lengthy liquidation process, unpaid loans, collection agency fees and various legal charges. The unpaid loans and the recovery procedures contribute significantly to the rising cost and increasing length of the liquidation process. The DPFB does not have any decision support tools that can be use to guide on debt recovery for institutions in liquidation. They rely on manual methods such us response to demand notes, the existence of security and the general availability of documents on the loan to classify whether a loan is a potentially good or bad debt.

Institutions in liquidation have massive loans data, which have been utilised in this project to aid the learning process of data mining tools in evaluating whether a particular debtor is likely to pay their debts. These results are meant to act as an enhancement to the mentioned manual methods.

This research explored the effectiveness of various data mining tools in evaluating whether a debt is likely to be repaid. The research involved recognition of the manual methods DPFB uses in classifying debts, the IT measures taken by DPFB to enhance debt recovery and the level of success achieved so far.

Loans data was collected from 27 institutions' databases. Data was prepared by selecting suitable variables (predictive and target). The predictive variables were four, namely; amount at liquidation, contacts availability, debt type and customer type. The target variable was the indication on whether a debt is good or bad.

Seven data mining methods were selected based on guidance from literature review. Predictive modelling software – DTREG was employed to train and validate the data mining techniques. The performance of each of the model was analysed using confusion matrix and area under receiver operating curve. Results were obtained from both balanced and imbalanced data. Balanced data performed better than imbalanced. The results which are detailed in chapter four depict that neural networks tools generally gave high accuracy. These findings guided to the development of a DSS prototype based on neural networks. This model can be used to aid DPFB decision support on debt recovery.

A neural network based DSS software - Alyuda Forecaster XL whose details a highlighted in section 4.5 formed the Model Based Management System (MBMS) component of the developed DSS prototype. On testing the software using different number of debt records, it was noted its classification accuracy increased with higher number of records.

The Data Base Management System (DBMS) part of the DSS was based on SQL 2000 while the Dialog Generation and Management System (DGMS) was based on Visual Studio and Crystal Report software.

In conjunction with the manual methods currently used by DPFB to classify debts, the developed decision support system is expected to enhance the accuracy of classification that may eventually lead to curbing the lengthy liquidation period and reduce expenses incurred in the process.

ACKNOWLEDGEMENTS

I wish to express my profound gratitude first and foremost to my God creator of heaven and earth who gave me the insight, ability and strength to be able to conceptualize and perform this study. Many thanks also goes to my supervisor in this project, Mr. Lawrence Muchemi, who is a lecturer at the School of Computing and Informatics, University of Nairobi for his kind and tireless guidance throughout the research. In particular he helped me moderate my ideas and ensured that the study was on track and met the required masters' level standards.

My thanks are also extended to my dear wife Grace for her constant encouragement and support throughout the study. To my children Victor, Joy, Praises and Enoch whose comments and mere presence around me was a great encouragement.

To my workmates in DPFb who enhanced my understanding of the debt recovery processes especially the manual methods used in classifying debts.

TABLE OF CONTENTS

DECLARATION.....	i
ABSTRACT	ii
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	iv
LIST OF FIGURES	vii
LIST OF TABLES	viii
LIST OF TERMINOLOGY AND ABBREVIATIONS.....	ix
CHAPTER ONE.....	1
INTRODUCTION	1
1.1 Background of the study	1
1.2 Statement of the problem.....	1
1.3 Purpose of the study.....	2
1.4 Objective of the study	2
1.5 Research questions.....	3
1.6 Significance of the study.....	3
1.7 Limitations of the study	4
1.8 Assumption of the study	4
1.9 Organization of the study.....	4
CHAPTER TWO.....	6
LITERATURE REVIEW	6
2.0 Introduction	6
2.1 Contributions to bad debt recovery by scholars.....	6
2.2 CHARACTERISTICS OF SAMPLED DATAMINING TOOLS	9
2.2.1 K- Nearest Neighbor (K-NN) - (Case- or Memory-based Reasoning)	9
Radial Basis Function	17
2.2.6 K-Means Clustering	23
2.2.7 Logistic Regression.....	26
2.2.8 Knowledge discovery and Data Mining.....	28
2.3 Decision Support Systems	29
2.3.1 Introduction.....	29
2.3.2 Characteristics of Decision Support Systems.....	29
2.3.3 Examples of DSS Solutions and their Utilization	30
2.3.5 Architecture of decision support systems.....	33
CHAPTER THREE.....	36
RESEARCH AND METHODOLOGY	36
3.1 Introduction	36

3.3 Location of the study	37
3.4 Target population	37
3.5 Data Mining Methodology.....	38
3.6 Study variables.....	38
3.7 Data Analysis and Validation Techniques	39
3.7.1. Simple validation	39
3.7.2 Cross validation	39
3.7.3. Confusion Matrix	40
3.7.4 Area under ROC curve.....	40
3.7.5. Prototype Design and Development.....	41
CHAPTER FOUR	42
ANALYSIS OF RESULTS AND DISCUSSIONS	42
4.1 Project Data	42
4.1.0 Data Collection	42
4.1.1 Data Preparation.....	43
4.1.2 Selected Data Mining Tools.....	44
4.2 Training and Validation	44
4.2.1 Cross Validation.....	44
4.2.2 Confusion Matrix	44
4.2.3. Area under ROC curve.....	45
4.3 Results of Models Assessment Based on Confusion Matrix & ROC Charts	45
4.3.1 Decision Tree	45
4.3.2 Neural Networks Multilayer Perception with 4 layers (2 hidden).....	47
4.3.3 Neural Networks Multilayer Perceptron with 3 layers(1 hidden)	50
4.3.4 Radial Basis Function (RBF) Neural Networks – Conceptually similar to K-	
4.3.5 Probabilistic Neural Networks (PNN) and General Regression Neural Networks (GRNN) –....	54
4.3.7 K-Mean Clustering.....	60
4.4 Design of the Prototype	63
4.4.1 Input Table Design.....	66
4.4.2 Output Design	66
4.4.5 Program Design	69
4.5.1 Testing of Alyuda Forecaster XL.....	70
4.5.2 Size of Data and Accuracy in Alyuda Forecaster XL	71
CHAPTER FIVE	72
5.0 CONCLUSIONS AND RECOMMENDATIONS	72
5.1 REVIEW OF OBJECTIVES OF THE STUDY	72
5.2 ACHIEVEMENTS OF THE STUDY	72

5.2.1 IT Measures Taken to Enhance Debt Recovery	72
5.2.2 Methods used to Classify Debts	72
base their classification of good and bad debts on:	72
5.2.3 Success Achieved in Debt Recovery	73
5.2.4 Determination of Appropriate Data Mining Tool	73
5.2.5 Development of DSS Prototype	73
5.3 CONCLUSIONS	74
5.3 RECOMMENDATIONS.....	75
5.4 SUGGESTIONS FOR FURTHER WORK	75
REFERENCES	76
APPENDICES	78
APPENDIX I: 4.6 A DECISION SUPPORT SYSTEM PROTOTYPE	78
APPENDIX II: SAMPLE DATA AND OUTPUT.....	96
APPENDIX III: PROTOTYPE SAMPLE CODE SAMPLE	98
APPENDIX IV: LETTER OF INTRODUCTION TO THE DIRECTOR – DPFb	101
APPENDIX V: REPLY LETTER FROM THE DIRECTOR - DPFb	102
APPENDIX VI: GUIDE TO USING DT-REG DEMONSTRATION SOFTWARE	103
.....	105

