

Decision support system for bandwidth expansion according to the increment of users

U J Hashan Dinuka

Index No-169311G

Faculty of Information Technology

University of Moratuwa

2019

Decision support system for bandwidth expansion according to the increment of users

U J Hashan Dinuka

Index No-169311G

Dissertation submitted to the Faculty of Information Technology,
University of Moratuwa, Sri Lanka for the partial fulfillment of the
Requirements of the Honors Degree of Bachelor of Science in
Information Technology.

February 2019

Declaration

We declare that is our own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Name of Student

Signature of Student

U J Hashan Dinuka

.....

Date:

Supervised by

Name of Supervisor

Signature of Supervisor

Mr. Saminda Premaratne

.....

Date:

Acknowledgements

First and foremost, it would be my pleasure to manifest heartfelt gratitude to the supervisor Mr. Saminda Premarathne, Senior Lecturer, Faculty of Information Technology, University of Moratuwa for his guidance without whom, this thesis would not have been possible. I should emphasize the thorough instructions from Prof. Asoka S. Karunananda helped me a lot. Also I would like to convey my sincere gratitude with all the professors who taught us throughout the M.Sc program for molding me to make this thesis a success.

I would also share my gratitude with the Lanka Bell (pvt.) Ltd. for the constant assistance throughout this process in terms of obtaining data. I would also like to thank all the batch mates of the M.Sc. in IT degree program who supported me with guidance and with their genuine and valuable opinions which led me to make the necessary improvements and finalized the results of the research and my family for the support they provided me.

Abstract

Considering the local context, in Sri Lanka internet speediness has become a strain issue domestically and in organization status making the promising promises of internet providers of delivering the fastest internet speed, an abortive attempt. One of the major reasons to that is giving a fixed bandwidth to the area, or fixed number of ENodeB s (EUTRAN Node B / Evolved Node B in 4g LTE). But with the increase of usage that given bandwidth becomes insufficient causing a lot of network traffic, mostly in peak times it almost becomes unusable as same as with ENodeBs becomes insufficient because too many clients connecting to the same node. The thorough analysis about this particular circumstance leads to understand that this issue has been occurred in different countries and most of the expertise has looked upon the matter. With the understanding that bandwidth expansion according to the user increment is a quite common concern which has been undergone with several suggested solutions; this research is going to explore the matter in terms of the local context considering the sample of consumer usage data of Lankabell (pvt) Ltd. The collected data will be processed through data mining algorithms. The objectives of this paper are to identify the internet speed decrements, high-risk area and low-risk area customers by using the data mining techniques like clustering.

KEYWORDS - Data mining, Clustering, Internet speed, Bandwidth, Broadband speed

Table of Contents

Declaration	i
Acknowledgements	ii
Abstract	iii
List of Tables	vii
List of Figures	vii
Chapter 01	1
Introduction.....	1
1.1 Prolegomenon.....	1
1.2 Research Aim & Objective.....	3
1.3 Problem Statement.....	3
1.4 Scope of the research.....	4
1.5 Proposed Solution.....	4
1.6 Summery.....	5
1.6.1 Structure of the Thesis.....	5
1.6.2 Chapter summery	5
Chapter 2.....	6
Review of data mining techniques and bandwidth expansion in substantive findings.....	6
2.1 Introduction	6
2.2 Related findings with reference to Data mining.....	6
2.3 Related findings with reference to Bandwidth Expansion requirement.....	7
2.4 Post Script.....	17
2.5 Summary.....	18
Chapter 03.....	19
Methodology	19
3.1 Introduction to the chapter.....	19
3.2 Technologies used for the Tool	19
3.2.1 What is Data mining.....	19
3.2.2 Cross-industry standard process.....	21
3.2.3 Data mining tool- RapidMiner	22

3.2.4 Java/ JSF/HTML	23
3.2.5 QoS Attribute	23
3.2.6 Average cause for Record closing (AVG_CAUSE_FOR_REC_CLOSING) ..	24
3.2.7 MS SQL.....	24
3.2.8 MS Excel	24
3.3 Summery	25
Chapter 04.....	26
Novel approach to predict the Bandwidth where the infrastructure upgrade is need the most according to the user/ usage	26
4.1 Introduction	26
4.2 Hypothesis	26
4.3 Process Model.....	27
4.4 System Over view.....	29
4.5 Raw Data Presentation & Inputs	30
4.5.1 Inputs	30
4.6 Output	31
4.7 Stakeholders of the solution	31
4.8 Features of the solution.....	31
4.9 Summery.....	32
Chapter 05	33
Implementation	33
5.1 Introduction	33
5.2 Top Level Architecture Design (System Design) of the solution	33
5.3 Data Collection and Data Preprocessing	34
5.3.1 Data Collection.....	34
5.3.2 Data Preprocessing.....	36
5.3.3 Data Cleaning.....	37
5.4 User Interface Model	38
5.5 Rapidminer Analytic Process	38
5.6 Back end Database.....	39
5.7 Summery.....	39

Chapter 6.....	40
Discussion	40
6.1 Introduction	40
6.2 Overall Implementation	40
6.3 Implementation of Rapidminer Analytical Engine.....	41
6.4 Summery.....	44
Chapter 07	45
Evaluation	45
7.1 Introduction	45
7.2 Evaluation of Classification techniques	45
7.3 Evaluation of the solution.....	46
7.4 Summery.....	50
Chapter 08.....	51
Conclusion and further work	51
8.1 Introduction	51
8.2 Limitation	52
8.3 Further Development.....	52
8.4 Summery.....	53
Bibliography	54
Appendixes - A	55
Appendixes - B	56
Appendixes - C	57
Appendixes - D	58

List of Tables

Table 1 Attributes (Inputs)	30
Table 2 Classification Evaluation Measuremets.....	45
Table 3 Comparison of classification techniques to determine overall outcome	46

List of Figures

Figure 1 Data Mining Process.....	20
Figure 2 Data Mining Paradigm	21
Figure 3 System Overview.....	29
Figure 4 Raw Data Sample	30
Figure 5 System Design.....	33
Figure 6 Raw Data Collection.....	34
Figure 7 Raw Data Status	35
Figure 8 Data Cleaning (Pre-preparation).....	35
Figure 9 Assortment of Data.....	37
Figure 10 Data assorting Process.....	37
Figure 11 Rapidminer analytic process.....	39
Figure 12 Dashboard Statistics	41
Figure 13 Implementation of analytic process.....	42
Figure 14 Decision Tree structure.....	43
Figure 15 Decision Tree Predictions.....	43
Figure 16 Naive Bayes Predictions.....	43
Figure 17 K-NN Predictions	44
Figure 18 Predicted feedback of Users	47
Figure 19 Original feedback from users	47

Figure 20 Predicted results of Galle Data set.....	48
Figure 21 Actual results of Galle Data set	48
Figure 22 Predicted data set - Alutgama.....	49
Figure 23 Actual feedback for January (Result)	49