FOREWORD

Praise be unto the almighty God for only by His grace and guidance that the author is able to finish this thesis titled "**A PERFORMANCE COMPARISON OF LOAD BALANCER ALGORITHM ON VIRTUALIZED ENVIRONMENT: CASE STUDY ON ROUND ROBIN, LEAST CONNECTION, AND WEIGHTED LEAST CONNECTION USING MACHINE LEARNING**" which is meant to fulfill a part of academic requirements to obtain the degree of Sarjana Komputer in Universitas Pelita Harapan, Tangerang.

The completion of this thesis would never have been possible without the support, guidance, and contribution from various parties involved. Therefore, the author would like to express utmost gratitude towards the following parties:

1) Dr.Eng., Ir. Pujianto Yugopuspito, M.Sc., as Dean of School of Information Science and Technology, Universitas Pelita Harapan and also the author's Thesis Co-Advisor, who has provided the author with many insights and suggestions to make this thesis better.
2) Ms. Irene Astuti Lazarusli, S.Kom., M.T., as Department Chair of Informatics, Universitas Pelita Harapan, who has provided the author with many useful and important knowledge throughout the author's academic journey.

3) Mr. I Made Murwantara, S.Si., M.Kom., Ph.D., as the author's Thesis Advisor, who has provided the author with invaluable guidance and constant support throughout the process of writing this thesis.
4) Mr. Frans Panduwinata, S.Kom., M.T., as the author's Academic Advisor, who has given constant motivation, guidance, and inspiration throughout the author's academic journey.
5) Lecturers and staffs in Informatics Department, Universitas Pelita Harapan, who have shared their knowledge to the author throughout the author's academic journey.
6) The author's mother and late father, for their never ending prayers and encouragement throughout the author's life.
7) Kezia Yunandra Kurniawan, who has given plenty emotional support and motivation throughout the process of writing this thesis.
8) The author's closest friends: Ryo Hansel Andersen, Aaron Maden Wilson, Michelle Evalista Patricia, William Pranata, Sandy Jaya Kusuma, Aaron Davis Sentono, Veronica Eviona, Natanaella Prayudhan, Graciela Dwi Cahya, Putri Larasati, Vanessa Evlin, and Aulia Kharisma Putri for their constant support and encouragement.
9) All members of Himpunan Mahasiwa Program Studi Informatika UPH (HMPTIF UPH) 2020/2021, in which the author enjoyed working together as a team to help create a warm and welcoming community for UPH Informatics students.

10) The big family of Mentoring UPH 2019/2020, especially to the author's Chief Mentor, Rahel Kristhea, who has given precious advice and motivation throughout the author's journey as a Mentor.
11) The UPH Informatics community from cohorts 2017, 2018, 2019, and 2020, who have supported the author in many ways.
12) All other parties who have helped the author either directly or indirectly in the completion of this thesis.

Finally, the author wishes that this thesis can be beneficial for all readers to broaden their knowledge, and also be used as a supporting material for further research, resulting in greater benefits.

Tangerang, 27 September 2021

Richard David Tedja

TABLE OF CONTENTS

TITLE PAGE ..... **i**
FINAL ASSIGNMENT STATEMENT AND UPLOAD AGREEMENT ..... **ii**
PERSETUJUAN DOSEN PEMBIMBING TUGAS AKHIR ..... **iii**
PERSETUJUAN TIM PENGUJI TUGAS AKHIR ..... **iv**
ABSTRACT ..... **v**
ABSTRAK ..... **vi**
FOREWORD ..... **vii**
TABLE OF CONTENTS ..... **x**
TABLE OF FIGURES ..... **xiii**
TABLE OF TABLES ..... **xix**
CHAPTER I INTRODUCTION **1**
1.1 Background 1
1.2 Problem Formulation 3
1.3 Scope Limitations 3
1.4 Purpose 4
1.5 Methodology 4
CHAPTER II LITERATURE REVIEW **5**
2.1 Load Balancing 5
2.2 Virtualization 6
2.3 Workload 8
2.4 Round Robin Algorithm 9
2.5 Least Connection Algorithm 10
2.6 Weighted Least Connection Algorithm 11
2.7 Receiver Operating Characteristics Analysis 11
2.8 Confusion Matrix 13
CHAPTER III RESEARCH METHODOLOGY **15**
3.1 Research Workflow 15
3.2 Research Design 16
3.2.1 Virtualized Environment 16

3.2.2 Load Balancer 17
3.2.3 Workload 19
3.3 Experiment Scenario 21
3.4 Data Collection 23
3.5 Data Analysis 27
CHAPTER IV EXPERIMENT RESULTS 29
4.1 Test Plan 1 Workload Testing Results 29
4.1.1 Test Plan 1 with Round Robin Algorithm 29
4.1.2 Test Plan 1 with Least Connection Algorithm 31
4.1.3 Test Plan 1 with Weighted Least Connection Algorithm 32
4.1.4 Algorithm Performance Comparison on Test Plan 1 36
4.2 Test Plan 2 Workload Testing Results 38
4.2.1 Test Plan 2 with Round Robin Algorithm 38
4.2.2 Test Plan 2 with Least Connection Algorithm 42
4.2.3 Test Plan 2 with Weighted Least Connection Algorithm 46
4.2.4 Algorithm Performance Comparison on Test Plan 2 53
4.3 Test Plan 3 Workload Testing Results 62
4.3.1 Test Plan 3 with Round Robin Algorithm 62
4.3.2 Test Plan 3 with Least Connection Algorithm 66
4.3.3 Test Plan 3 with Weighted Least Connection Algorithm 70
4.3.4 Algorithm Performance Comparison on Test Plan 3 77
4.4 Test Plan 4 Workload Testing Results 87
4.4.1 Test Plan 4 with Round Robin Algorithm 87
4.4.2 Test Plan 4 with Least Connection Algorithm 91
4.4.3 Test Plan 4 with Weighted Least Connection Algorithm 95
4.4.4 Algorithm Performance Comparison on Test Plan 4 101
4.5 Test Plan 5 Workload Testing Results 111
4.5.1 Test Plan 5 with Round Robin Algorithm 111
4.5.2 Test Plan 5 with Least Connection Algorithm 115
4.5.3 Test Plan 5 with Weighted Least Connection Algorithm 119
4.5.4 Algorithm Performance Comparison on Test Plan 5 125
4.6 Test Plan 6 Workload Testing Results 135

4.6.1 Test Plan 6 with Round Robin Algorithm 135
4.6.2 Test Plan 6 with Least Connection Algorithm 139
4.6.3 Test Plan 6 with Weighted Least Connection Algorithm 143
4.6.4 Algorithm Performance Comparison on Test Plan 6 149
CHAPTER V ANALYSIS 159
5.1 Datasets 159
5.1.1 Train Dataset 159
5.1.2 Test Dataset 160
5.2 Machine Learning Model 162
5.2.1 Model Configuration 162
5.2.2 Model Evaluation 163
5.3 Prediction Output 168
CHAPTER VI CONCLUSIONS AND SUGGESTIONS 170
6.1 Conclusions 170
6.2 Suggestions 171
REFERENCES 173

TABLE OF FIGURES

Figure 2.1 Scheduling Model in the Cloud Data Center 5
Figure 2.2 Scenario of Operating System-based Virtualization 7
Figure 2.3 Illustration of Round Robin Algorithm 10
Figure 2.4 Load Balancing with Least Connection Algorithm 11
Figure 2.5 ROC Graph with Three ROC Curves 12
Figure 2.6 Multi-class Classification Confusion Matrix 14
Figure 3.1 Research Workflow 16
Figure 3.2 Virtualized Environment Instances 17
Figure 3.3 HAProxy Configuration File with Algorithms Implemented 18
Figure 3.4 HAProxy Statistics Monitoring Page 19
Figure 3.5 Experiment Scenario Flowchart 22
Figure 3.6 Ansible Playbook 23
Figure 3.7 Data Collection Flowchart 24
Figure 3.8 Bash Script for Automatic Data Collection 27
Figure 4.1 Algorithm Frontend Performance Comparison on Test Plan 1 37
Figure 4.2 Algorithm Backend Performance Comparison on Test Plan 1 37
Figure 4.3 Test Plan 2 Round Robin Frontend Performance (Threads Sent/Processed) 39
Figure 4.4 Test Plan 2 Round Robin Frontend Performance (Session Send/Process Rate) 40
Figure 4.5 Test Plan 2 Round Robin Backend Performance (Threads Received) 41
Figure 4.6 Test Plan 2 Round Robin Backend Performance (Session Receive Rate) 42
Figure 4.7 Test Plan 2 Least Connection Frontend Performance (Threads Sent/Processed) 43
Figure 4.8 Test Plan 2 Least Connection Frontend Performance (Session Send/Process Rate) 44
Figure 4.9 Test Plan 2 Least Connection Backend Performance (Threads Received) 45
Figure 4.10 Test Plan 2 Least Connection Backend Performance (Session Receive Rate) 46
Figure 4.11 Test Plan 2 Weighted Least Connection Frontend Performance (Threads Sent) 47
Figure 4.12 Test Plan 2 Weighted Least Connection Frontend Performance (Session Send Rate) 48
Figure 4.13 Test Plan 2 Weighted Least Connection Frontend Performance (Average Session Send Rate) 48
Figure 4.14 Test Plan 2 Weighted Least Connection Frontend Performance (Threads Processed) 49

Figure 4.15 Test Plan 2 Weighted Least Connection Frontend Performance (Session Process Rate) 50
Figure 4.16 Test Plan 2 Weighted Least Connection Frontend Performance (Average Session Process Rate) 50
Figure 4.17 Test Plan 2 Weighted Least Connection Frontend Performance (Threads Received) 51
Figure 4.18 Test Plan 2 Weighted Least Connection Backend Performance (Session Receive Rate) 52
Figure 4.19 Test Plan 2 Weighted Least Connection Backend Performance (Average Session Receive Rate) 52
Figure 4.20 Algorithm Frontend Performance Comparison on Test Plan 2 (Threads Sent) 53
Figure 4.21 Algorithm Frontend Performance Comparison on Test Plan 2 (Session Send Rate) 54
Figure 4.22 Algorithm Frontend Performance Comparison on Test Plan 2 (Average Session Send Rate) 55
Figure 4.23 Algorithm Frontend Performance Comparison on Test Plan 2 (Threads Processed) 55
Figure 4.24 Algorithm Frontend Performance Comparison on Test Plan 2 (Session Process Rate) 56
Figure 4.25 Algorithm Frontend Performance Comparison on Test Plan 2 (Average Session Process Rate) 57
Figure 4.26 Algorithm Frontend Performance Comparison on Test Plan 2 (Threads Received) 58
Figure 4.27 Figure 4. 23 Algorithm Frontend Performance Comparison on Test Plan 2 (Session Receive Rate) 61
Figure 4.28 Algorithm Backend Performance Comparison on Test Plan 2 (Average Session Receive Rate) 61
Figure 4.29 Test Plan 3 Round Robin Frontend Performance (Threads Sent/Processed) 63
Figure 4.30 Test Plan 3 Round Robin Frontend Performance (Session Send/Process Rate) 64
Figure 4.31 Test Plan 3 Round Robin Backend Performance (Threads Received) 65
Figure 4.32 Test Plan 3 Round Robin Backend Performance (Session Receive Rate) 66
Figure 4.33 Test Plan 3 Least Connection Frontend Performance (Threads Sent/Processed) 67
Figure 4.34 Test Plan 3 Least Connection Frontend Performance (Session Rate) 68
Figure 4.35 Test Plan 3 Least Connection Backend Performance (Threads Received) 69

Figure 4.36 Test Plan 3 Least Connection Backend Performance (Session Receive Rate) 70
Figure 4.37 Test Plan 3 Weighted Least Connection Frontend Performance (Threads Sent) 71
Figure 4.38 Test Plan 3 Weighted Least Connection Frontend Performance (Session Send Rate) 72
Figure 4.39 Test Plan 3 Weighted Least Connection Frontend Performance (Session Send Rate) 72
Figure 4.40 Test Plan 3 Weighted Least Connection Frontend Performance (Threads Processed) 73
Figure 4.41 Test Plan 3 Weighted Least Connection Frontend Performance (Session Process Rate) 74
Figure 4.42 Test Plan 3 Weighted Least Connection Frontend Performance (Average Session Process Rate) 74
Figure 4.43 Test Plan 3 Weighted Least Connection Frontend Performance (Threads Received) 75
Figure 4.44 Test Plan 3 Weighted Least Connection Backend Performance (Session Receive Rate) 76
Figure 4.45 Test Plan 3 Weighted Least Connection Backend Performance (Average Session Receive Rate) 77
Figure 4.46 Algorithm Frontend Performance Comparison on Test Plan 3 (Threads Sent) 78
Figure 4.47 Algorithm Frontend Performance Comparison on Test Plan 3 (Session Send Rate) 78
Figure 4.48 Algorithm Frontend Performance Comparison on Test Plan 2 (Average Session Sent Rate) 79
Figure 4.49 Algorithm Frontend Performance Comparison on Test Plan 3 (Threads Processed) 80
Figure 4.50 Algorithm Frontend Performance Comparison on Test Plan 3 (Session Process Rate) 80
Figure 4.51 Algorithm Frontend Performance Comparison on Test Plan 3 (Average Session Process Rate) 82
Figure 4.52 Algorithm Frontend Performance Comparison on Test Plan 3 (Threads Received) 84
Figure 4.53 Algorithm Frontend Performance Comparison on Test Plan 3 (Session Receive Rate) 86
Figure 4.54 Algorithm Backend Performance Comparison on Test Plan 3 (Average Session Receive Rate) 86
Figure 4.55 Test Plan 4 Round Robin Frontend Performance (Threads Sent/Processed) 88
Figure 4.56 Test Plan 4 Round Robin Frontend Performance (Session Send/Process Rate) 88

Figure 4.57 Test Plan 4 Round Robin Backend Performance (Threads Received) 89
Figure 4.58 Test Plan 4 Round Robin Backend Performance (Session Receive Rate) 90
Figure 4.59 Test Plan 4 Least Connection Frontend Performance (Threads Sent/Processed) 92
Figure 4.60 Test Plan 4 Least Connection Frontend Performance (Session Send/Process Rate) 92
Figure 4.61 Test Plan 4 Least Connection Backend Performance (Threads Received) 94
Figure 4.62 Test Plan 4 Least Connection Backend Performance (Session Receive Rate) 94
Figure 4.63 Test Plan 4 Weighted Least Connection Frontend Performance (Threads Sent) 96
Figure 4.64 Test Plan 4 Weighted Least Connection Frontend Performance (Session Send Rate) 96
Figure 4.65 Test Plan 4 Weighted Least Connection Frontend Performance (Average Session Send Rate) 97
Figure 4.66 Test Plan 4 Weighted Least Connection Frontend Performance (Threads Processed) 98
Figure 4.67 Test Plan 4 Weighted Least Connection Frontend Performance (Session Process Rate) 98
Figure 4.68 Test Plan 4 Weighted Least Connection Frontend Performance (Average Session Process Rate) 99
Figure 4.69 Test Plan 4 Weighted Least Connection Frontend Performance (Threads Received) 100
Figure 4.70 Test Plan 4 Weighted Least Connection Backend Performance (Session Receive Rate) 100
Figure 4.71 Test Plan 4 Weighted Least Connection Backend Performance (Average Session Receive Rate) 101
Figure 4.72 Algorithm Frontend Performance Comparison on Test Plan 4 (Threads Sent) 102
Figure 4.73 Algorithm Frontend Performance Comparison on Test Plan 4 (Session Send Rate) 103
Figure 4.74 Algorithm Frontend Performance Comparison on Test Plan 4 (Average Session Send Rate) 103
Figure 4.75 Algorithm Frontend Performance Comparison on Test Plan 4 (Threads Processed) 104
Figure 4.76 Algorithm Frontend Performance Comparison on Test Plan 4 (Session Process Rate) 105
Figure 4.77 Algorithm Frontend Performance Comparison on Test Plan 4 (Average Session Process Rate) 106

Figure 4.78 Algorithm Frontend Performance Comparison on Test Plan 4 (Threads Received) 108
Figure 4.79 Algorithm Frontend Performance Comparison on Test Plan 4 (Session Receive Rate)..... 110
Figure 4.80 Algorithm Backend Performance Comparison on Test Plan 4 (Average Session Receive Rate) 110
Figure 4.81 Test Plan 5 Round Robin Frontend Performance (Threads Sent/Received) 112
Figure 4.82 Test Plan 5 Round Robin Frontend Performance (Session Send/Process Rate) 112
Figure 4.83 Test Plan 5 Round Robin Backend Performance (Threads Received) 113
Figure 4.84 Test Plan 5 Round Robin Backend Performance (Session Receive Rate)..... 114
Figure 4.85 Test Plan 5 Least Connection Frontend Performance (Threads Sent/Received) 116
Figure 4.86 Test Plan 5 Least Connection Frontend Performance (Session Send/Process Rate) 116
Figure 4.87 Test Plan 5 Least Connection Backend Performance (Threads Received) 118
Figure 4.88 Test Plan 5 Least Connection Backend Performance (Session Receive Rate)..... 118
Figure 4.89 Test Plan 5 Weighted Least Connection Frontend Performance (Threads Sent) 120
Figure 4.90 Test Plan 5 Weighted Least Connection Frontend Performance (Session Send Rate) 120
Figure 4.91 Test Plan 5 Weighted Least Connection Frontend Performance (Session Send Rate) 121
Figure 4.92 Test Plan 5 Weighted Least Connection Frontend Performance (Threads Processed) 122
Figure 4.93 Test Plan 5 Weighted Least Connection Frontend Performance (Session Process Rate) 122
Figure 4.94 Test Plan 5 Weighted Least Connection Frontend Performance (Average Session Process Rate)..... 123
Figure 4.95 Test Plan 5 Weighted Least Connection Frontend Performance (Threads Received) 124
Figure 4.96 Test Plan 5 Weighted Least Connection Backend Performance (Session Receive Rate)..... 124
Figure 4.97 Test Plan 5 Weighted Least Connection Backend Performance (Average Session Receive Rate) 125
Figure 4.98 Algorithm Frontend Performance Comparison on Test Plan 5 (Threads Sent) 126

Figure 4.99 Algorithm Frontend Performance Comparison on Test Plan 5 (Session Send Rate) 127
Figure 4.100 Algorithm Frontend Performance Comparison on Test Plan 5 (Average Session Send Rate) 127
Figure 4.101 Algorithm Frontend Performance Comparison on Test Plan 5 (Threads Processed) 128
Figure 4.102 Algorithm Frontend Performance Comparison on Test Plan 5 (Session Process Rate) 129
Figure 4.103 Algorithm Frontend Performance Comparison on Test Plan 5 (Average Session Process Rate) 130
Figure 4.104 Algorithm Frontend Performance Comparison on Test Plan 5 (Threads Received) 131
Figure 4.105 Algorithm Frontend Performance Comparison on Test Plan 5 (Session Receive Rate) 134
Figure 4.106 Algorithm Backend Performance Comparison on Test Plan 5 (Average Session Receive Rate) 134
Figure 4.107 Test Plan 6 Round Robin Frontend Performance (Threads Sent/Processed) 136
Figure 4.108 Test Plan 6 Round Robin Frontend Performance (Session Send/Process Rate) 136
Figure 4.109 Test Plan 6 Round Robin Backend Performance (Threads Received) 137
Figure 4.110 Test Plan 6 Round Robin Backend Performance (Session Receive Rate) 138
Figure 4.111 Test Plan 6 Least Connection Frontend Performance (Threads Sent/Processed) 140
Figure 4.112 Test Plan 6 Least Connection Frontend Performance (Session Send/Process Rate) 140
Figure 4.113 Test Plan 6 Least Connection Backend Performance (Threads Received) 142
Figure 4.114 Test Plan 6 Least Connection Backend Performance (Session Receive Rate) 142
Figure 4.115 Test Plan 6 Weighted Least Connection Frontend Performance (Threads Sent) 144
Figure 4.116 Test Plan 6 Weighted Least Connection Frontend Performance (Session Send Rate) 144
Figure 4.117 Test Plan 6 Weighted Least Connection Frontend Performance (Session Send Rate) 145
Figure 4.118 Test Plan 6 Weighted Least Connection Frontend Performance (Threads Processed) 146
Figure 4.119 Test Plan 6 Weighted Least Connection Frontend Performance (Session Process Rate) 146

Figure 4.120 Test Plan 6 Weighted Least Connection Frontend Performance (Average Session Process Rate) 147
Figure 4.121 Test Plan 6 Weighted Least Connection Frontend Performance (Threads Received) 148
Figure 4.122 Test Plan 6 Weighted Least Connection Backend Performance (Session Receive Rate) 148
Figure 4.123 Test Plan 6 Weighted Least Connection Backend Performance (Average Session Receive Rate) 149
Figure 4.124 Algorithm Frontend Performance Comparison on Test Plan 6 (Threads Sent) 150
Figure 4.125 Algorithm Frontend Performance Comparison on Test Plan 6 (Session Send Rate) 151
Figure 4.126 Algorithm Frontend Performance Comparison on Test Plan 6 (Average Session Send Rate) 151
Figure 4.127 Algorithm Frontend Performance Comparison on Test Plan 6 (Threads Processed) 152
Figure 4.128 Algorithm Frontend Performance Comparison on Test Plan 6 (Session Process Rate) 153
Figure 4.129 Algorithm Frontend Performance Comparison on Test Plan 6 (Average Session Process Rate) 154
Figure 4.130 Algorithm Frontend Performance Comparison on Test Plan 6 (Threads Received) 155
Figure 4.131 Algorithm Frontend Performance Comparison on Test Plan 6 (Session Receive Rate) 158
Figure 4.132 Algorithm Backend Performance Comparison on Test Plan 6 (Average Session Receive Rate) 158
Figure 5.1 Confusion Matrix for AdaBoost 165
Figure 5.2 Confusion Matrix for kNN 165
Figure 5.3 Confusion Matrix for Random Forest 165
Figure 5.4 ROC Analysis for Target Class RR 167
Figure 5.5 ROC Analysis for Target Class LC 167
Figure 5.6 ROC Analysis for Target Class WLC 168

TABLE OF TABLES

Table 3.1 Planned Test Plans 21
Table 3.2 Sample Cleansed HAProxy Metrics Data 21
Table 3.3 Sample Cleansed JMeter Test Report Data 21
Table 4.1 Test Plan 1 with Round Robin Algorithm Frontend Results 30
Table 4.2 Test Plan 1 with Round Robin Algorithm Backend Results 30
Table 4.3 Test Plan 1 with Least Connection Algorithm Frontend Results 31
Table 4.4 Least Connection Algorithm Backend Results 32
Table 4.5 Test Plan 1 with Weighted Least Connection Algorithm Frontend Results 33
Table 4.6 Test Plan 1 Weighted Least Connection Algorithm Backend Results 35
Table 4.7 Test Plan 2 Total Received Threads Comparison 59
Table 4.8 Test Plan 3 Total Received Threads Comparison 83
Table 4.9 Test Plan 4 Total Received Threads Comparison 107
Table 4.10 Test Plan 5 Total Received Threads Comparison 132
Table 4.11 Test Plan 6 Total Received Threads Comparison 156
Table 5.1 First 10 Instances of Train Dataset 160
Table 5.2 Test Dataset Sample for Predictions 161
Table 5.3 Algorithm Performance Measurement Scores 163
Table 5.4 Prediction Output 169