



UNDERGRADUATE THESIS

COMPARISON OF XGBOOST AND LIGHTGBM METHODS WITH RANDOM OVERSAMPLING FOR OBESITY LEVEL CLASSIFICATION

CHESA SASKIA RAFIKA
NPM 22081010211

THESIS ADVISORS

Eva Yulia Puspaningrum, S.Kom., M.Kom
Retno Mumpuni, S.Kom., M.Sc

**MINISTRY OF HIGHER EDUCATION, SCIENCE, AND TECHNOLOGY
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FACULTY OF COMPUTER SCIENCE
INFORMATICS STUDY PROGRAM
SURABAYA
2026**

APPROVAL SHEET

COMPARISON OF XGBOOST AND LIGHTGBM METHODS WITH RANDOM OVERSAMPLING FOR OBESITY LEVEL CLASSIFICATION

By:
CHESA SASKIA RAFIKA
NPM. 22081010211

Has been defended before, and accepted by, the Board of Assessors of the Thesis Examination of the Informatics Study Program, Faculty of Computer Science, Universitas Pembangunan Nasional Veteran Jawa Timur, on June 18, 2026:

Approved,

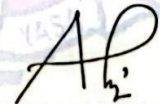
Eva Yulia Puspaningrum, S.Kom., M.Kom.
NIP. 19890705 202121 2 002


..... (Advisor I)

Retno Mumpuni, S.Kom., M.Sc.
NIP. 19870716 202521 2 045


..... (Advisor II)

Dr. Eng. Ir. Anggraeni Puspita Sari, ST., MT.
NIP. 222198 60 816400


..... (Head Assessor)

Ardhon Rakhmadi, S.Tr.T., M.Kom.
NIP. 19910805 202406 1 002


..... (Assessor I)

Acknowledge by,

Dean of the Faculty of Computer Science



Prof. Dr. Ir. Novirina Hendrasarie, MT.
NIP. 19681126 199403 2 001

APPROVAL SHEET

COMPARISON OF XGBOOST AND LIGHTGBM METHODS WITH RANDOM OVERSAMPLING FOR OBESITY LEVEL CLASSIFICATION

By:
CHESA SASKIA RAFIKA
NPM. 22081010211

Approved to proceed to the Thesis Examination



Approved by,

Coordinator of Informatics Study Program
Faculty of Computer Science

A handwritten signature in blue ink, belonging to Dr. Intan Yuniar Purbasari, is positioned above the printed name.

Dr. Intan Yuniar Purbasari, S.Kom., M.Sc.
NIP. 19800602 202521 2 029

STATEMENT OF ORIGINALITY

I am the undersigned:

Student Name : Chesa Saskia Rafika
NPM : 22081010211
Degree Program : Bachelor (S1)
Study Program : Informatics
Faculty : Faculty of Computer Science

Hereby declares that this undergraduate thesis contains no part of any other scientific work that has been submitted to obtain an academic degree at any higher education institution. Furthermore, it does not contain any work or opinions previously written or published by others, except for those which are explicitly cited in this thesis and listed completely in references.

And I declare that this scientific document is free from elements of plagiarism. If in the future indications of plagiarism are found in this Thesis, I am willing to accept sanctions in accordance with the applicable laws and regulations.

Thus, I made this statement without any coercion from anyone and to be used as it should.



Surabaya, July 6, 2026
Declarant,



CHESA SASKIA RAFIKA
NPM. 22081010211

ABSTRACT

Student Name / NPM : Chesa Saskia Rafika / 22081010211

Thesis Title : Comparison of XGBoost and LightGBM with Random Oversampling for Obesity Level Classification

Supervisors : 1. Eva Yulia Puspaningrum, S.Kom., M.Kom
2. Retno Mumpuni, S.Kom., M.Sc

Obesity is a health issue related to several aspects of individual condition and daily behavior, including physical characteristics, food consumption habits, physical activity, and lifestyle patterns. This study was conducted to evaluate the performance of XGBoost and LightGBM in classifying obesity levels and to analyze the use of Random Oversampling (ROS) for handling unequal class distribution. The research used the Dhaka Obesity Dataset, which initially contained 2,182 respondents with 17 attributes. The data preparation process involved selecting adult respondents by removing data under 18 years old, reconstructing the target labels based on the Asia Pacific Body Mass Index (BMI) classification adopted by the Indonesian Ministry of Health, cleaning the dataset, transforming categorical variables into numerical form, and dividing the data using stratified train test split. After preprocessing and duplicate removal, the final dataset consisted of 2,125 records. The experimental stage was carried out by testing several train test split ratios, applying stepwise manual hyperparameter tuning, and comparing model performance under baseline and ROS scenarios. Model evaluation was performed using accuracy, macro precision, macro recall, macro F1 score, balanced accuracy, confusion matrix, classification report, and train test gap. The results showed that ROS improved the performance of both XGBoost and LightGBM, particularly by producing more balanced classification results across obesity classes. Among all evaluated models, LightGBM with ROS achieved the highest performance, with an accuracy of 0.967059, macro precision of 0.955319, macro recall of 0.942773, macro F1 score of 0.948016, and balanced accuracy of 0.942773. Therefore, LightGBM with ROS was selected as the most suitable model for obesity level classification in this study.

Keywords: Obesity, XGBoost, LightGBM, Random Oversampling, Classification, Class Imbalance.

ACKNOWLEDGEMENTS

All praise and gratitude are directed to Allah SWT for His abundant blessings, guidance, and mercy, which have given the author the strength and opportunity to complete this undergraduate thesis entitled “Comparison of XGBoost and LightGBM with Random Oversampling for Obesity Level Classification.”

This thesis was written as a partial requirement for the completion of the Bachelor’s degree in the Informatics Study Program, Faculty of Computer Science, Universitas Pembangunan Nasional “Veteran” Jawa Timur. During the preparation and completion of this thesis, the author received valuable assistance, guidance, motivation, and prayers from many parties. The author realizes that this thesis could not have been completed properly without their support. Therefore, the author would like to convey sincere appreciation and gratitude to:

1. The author's beloved parents and family, for their endless love, prayers, encouragement, motivation, and unwavering support throughout the completion of this thesis.
2. Mrs. Eva Yulia Puspaningrum, S.Kom., M.Kom., as the First Supervisor, for her invaluable guidance, constructive suggestions, continuous support, and motivation throughout the research and writing process.
3. Mrs. Retno Mumpuni, S.Kom., M.Sc., as the Second Supervisor, for her guidance, insightful advice, and valuable recommendations that greatly contributed to the completion of this thesis.
4. Mrs. Dr. Eng. Ir. Anggraeni Puspita Sari, S.T., M.T., as the First Examiner, for her constructive criticism, valuable suggestions, and insightful feedback that helped improve the quality of this thesis.
5. Mr. Ardhon Rakhmadi, S.Tr.T., M.Kom., as the Second Examiner, for his valuable guidance, evaluation, and constructive feedback throughout the thesis examination process.
6. All lecturers and academic staff of the Informatics Study Program, Faculty of Computer Science, Universitas Pembangunan Nasional "Veteran" Jawa Timur, for their knowledge, guidance, experience, and support throughout the author's academic journey.

7. Dhea Angel Silfany, Vika Rafi Ana, and Revano Maliq Reynanda, for accompanying the author throughout university life, sharing both joyful and challenging moments, and continuously providing encouragement, assistance, and support during the completion of this thesis.
8. Friends and everyone else who cannot be mentioned individually, for their support, encouragement, and assistance throughout the preparation of this thesis.

The author realizes that this thesis still has limitations and may contain shortcomings. Therefore, constructive criticism and suggestions are sincerely welcomed for the improvement of this work in the future. The author hopes that this thesis can provide benefits to readers and serve as a useful reference for future research.

Surabaya, June 19, 2026

Author

TABLE OF CONTENTS

COVER PAGE.....	ii
APPROVAL SHEET.....	iii
APPROVAL SHEET.....	iv
STATEMENT OF ORIGINALITY	v
ABSTRACT.....;;.....	vi
ACKNOWLEDGEMENTS.....	vii
TABLE OF CONTENTS.....	ix
LIST OF FIGURES.....	xii
LIST OF TABLES.....	xviii
CHAPTER 1 INTRODUCTION.....	1
1.1. Research Background.....	1
1.2. Research Questions	4
1.3. Research Objectives	4
1.4. Research Significance	4
1.5. Scope and Limitations of the Study	5
CHAPTER II LITERATURE REVIEW	6
2.1. Previous Studies	7
2.2. Obesity	11
2.3. Factors Causing Obesity.....	13
2.4. Machine Learning	14
2.5. Supervised Learning.....	16
2.6. Classification.....	17
2.7. XGBoost.....	19
2.8. LightGBM	24
2.9. Random Oversampling.....	29

2.10. Model Evaluation and Confusion Matrix.....	31
CHAPTER III RESEARCH METHODOLOGY	36
3.1. Data Acquisition.....	37
3.2. Data Preprocessing.....	40
3.3. Dataset Adjustment Based on the Ministry of Health Standard.....	41
3.4. Data Cleaning.....	43
3.5. Encoding Data	47
3.6. Training and Testing Data Split	49
3.7. Class Imbalance Handling.....	50
3.8. Model Development.....	51
3.8.1. XGBoost	53
3.8.2. LightGBM	62
3.9. Model Evaluation	73
3.10. Model Analysis Scheme.....	75
3.11. Model Implementation Design.....	77
CHAPTER IV RESULTS AND DISCUSSION.....	82
4.1. Program Implementation.....	82
4.1.1. Library	82
4.1.2. Data Acquisition and Dataset Adjustment.....	83
4.1.3. Data Preprocessing	90
4.1.4. Data Balancing Using Random Oversampling.....	98
4.1.5. Model Development	99
4.1.6. Model Evaluation.....	100
4.2. Comparison of Testing Schemes.....	102
4.2.1. Testing of Data Split Ratios on XGBoost + ROS	104
4.2.2. Testing of Data Split Ratios on LightGBM + ROS.....	111
4.2.3. Learning Rate Testing.....	119

4.2.4. Max Depth Testing	132
4.2.5. n_estimators Testing	147
4.2.6. Subsample Testing.....	159
4.2.7. Colsample_bytree Testing	173
4.2.8. Final Evaluation of XGBoost + ROS and LightGBM + ROS.....	187
4.2.9. XGBoost and LightGBM without Random Oversampling	189
4.2.10. Analysis and Determination of the Best Model.....	194
4.3. System Interface Design Implementation Results.....	197
4.3.1. Login Page Implementation Results	197
4.3.2. Register Page Implementation Results	198
4.3.3. Home Page Implementation Results.....	199
4.3.4. Input Form Page Implementation Result.....	201
4.3.5. Result Page Implementation Results	202
CHAPTER V CONCLUSIONS AND RECOMMENDATION	210
5.1. Conclusions	210
5.2. Recommendations	211
REFERENCES.....	213
APPENDICES.....	218

LIST OF FIGURES

Figure 2.1. Body Mass Index (BMI) Categories.....	11
Figure 2.2. Categories of Machine Learning Algorithms	15
Figure 2.3. Supervised Learning.....	16
Figure 2.3. Supervised Learning.....	18
Figure 2.5. XGBoost Workflow.....	19
Figure 2.6. LightGBM Workflow.....	24
Figure 2.7. Random Oversampling	29
Figure 2.8. <i>Confusion Matrix</i>	32
Figure 3.1 Metodologi.....	36
Figure 3.2 Data Preprocessing.....	40
Figure 3.3. Random Oversampling (ROS) Workflow	51
Figure 3.4. XGBoost Model Development Workflow.....	53
Figure 3.5. LightGBM Model Development Workflow	62
Figure 3.6. Login Wireframe.....	78
Figure 3.7. Register Wireframe.....	79
Figure 3.8. Home Page Wireframe.....	79
Figure 3.9. Input Form Page Wireframe	80
Figure 3.10. Prediction Result Page Wireframe	81
Figure 4.1. Initial Dataset Information.....	84
Figure 4.2. Preview of the First Five Rows of the Dhaka Obesity Dataset	85
Figure 4.3. Examples of Five Respondents Under 18 Years of Age Removed.....	87
Figure 4.4. Number of Data Under 18 Years of Age Removed.....	87
Figure 4.5. Examples of Five Data Records After Target Relabeling	89
Figure 4.6. Changes in Target Class Distribution.....	90
Figure 4.7. Results of Column Name Standardization	92
Figure 4.9. Missing Value Checking Results.....	94

Figure 4.12. Train Test Graph of XGBoost + ROS Split 80:20	104
Figure 4.13. Confusion Matrix of XGBoost + ROS Split 80:20	105
Figure 4.14. Classification Report of XGBoost + ROS Split 80:20	105
Figure 4.15. Train Test Graph of XGBoost + ROS Split 70:30	106
Figure 4.16. Confusion Matrix of XGBoost + ROS Split 70:30	107
Figure 4.17. Classification Report of XGBoost + ROS Split 70:30	107
Figure 4.18. Train Test Graph of XGBoost + ROS Split 60:40	108
Figure 4.19. Confusion Matrix of XGBoost + ROS Split 60:40	109
Figure 4.20. Classification Report of XGBoost + ROS Split 60:40	109
Figure 4.21. Train Test Graph of LightGBM + ROS Split 80:20.....	112
Figure 4.22. Confusion Matrix of LightGBM + ROS Split 80:20.....	113
Figure 4.23. Classification Report of LightGBM + ROS Split 80:20	113
Figure 4.24. Train Test Graph of LightGBM + ROS Split 70:30.....	114
Figure 4.25. Confusion Matrix of LightGBM + ROS Split 70:30.....	115
Figure 4.26. Classification Report of LightGBM + ROS Split 70:30	115
Figure 4.27. Train Test Graph of LightGBM + ROS Split 60:40.....	116
Figure 4.28. Confusion Matrix of LightGBM + ROS Split 60:40.....	117
Figure 4.29. Classification Report of LightGBM + ROS Split 60:40	118
Figure 4.30. Train Test Graph of XGBoost + ROS learning_rate 0.01	120
Figure 4.31. Confusion Matrix of XGBoost + ROS learning_rate 0.01	120
Figure 4.32. Classification Report of XGBoost + ROS learning_rate 0.01.....	121
Figure 4.33. Train Test Graph of XGBoost + ROS learning_rate 0.05	122
Figure 4.34. Confusion Matrix of XGBoost + ROS learning_rate 0.05	122
Figure 4.35. Classification Report of XGBoost + ROS learning_rate 0.05.....	123
Figure 4.36. Train Test Graph of XGBoost + ROS learning_rate 0.10	123
Figure 4.37. Confusion Matrix of XGBoost + ROS learning_rate 0.10	124
Figure 4.38. Classification Report of XGBoost + ROS learning_rate 0.10.....	124

Figure 4.39. Train Test Graph of LightGBM + ROS learning_rate 0.01	125
Figure 4.40. Confusion Matrix of LightGBM + ROS learning_rate 0.01	126
Figure 4.41. Classification Report of LightGBM + ROS learning_rate 0.01	126
Figure 4.42. Train Test Graph of LightGBM + ROS learning_rate 0.05	127
Figure 4.43. Confusion Matrix of LightGBM + ROS learning_rate 0.05	128
Figure 4.44. Classification Report of LightGBM + ROS learning_rate 0.05	128
Figure 4.45. Train Test Graph of LightGBM + ROS learning_rate 0.10	129
Figure 4.46. Confusion Matrix of LightGBM + ROS learning_rate 0.10	130
Figure 4.47. Classification Report of LightGBM + ROS learning_rate 0.10	130
Figure 4.48. Train Test Graph of XGBoost + ROS max_depth 3	133
Figure 4.49. Confusion Matrix of XGBoost + ROS max_depth 3	134
Figure 4.50. Classification Report of XGBoost + ROS max_depth 3	134
Figure 4.51. Train Test Graph of XGBoost + ROS max_depth 5	135
Figure 4.52. Confusion Matrix of XGBoost + ROS max_depth 5	136
Figure 4.53. Classification Report of XGBoost + ROS max_depth 5	136
Figure 4.54. Train Test Graph of XGBoost + ROS max_depth 7	137
Figure 4.55. Confusion Matrix of XGBoost + ROS max_depth 7	138
Figure 4.56. Classification Report of XGBoost + ROS max_depth 7	138
Figure 4.57. Train Test Graph of LightGBM + ROS max_depth 3.....	139
Figure 4.58. Confusion Matrix of LightGBM + ROS max_depth 3.....	140
Figure 4.59. Classification Report of LightGBM + ROS max_depth 3	140
Figure 4.60. Train Test Graph of LightGBM + ROS max_depth 5.....	141
Figure 4.61. Confusion Matrix of LightGBM + ROS max_depth 5.....	142
Figure 4.62. Classification Report of LightGBM + ROS max_depth 5	142
Figure 4.63. Train Test Graph of LightGBM + ROS max_depth 7.....	143
Figure 4.64. Confusion Matrix of LightGBM + ROS max_depth 7.....	144
Figure 4.65. Classification Report of LightGBM + ROS max_depth 7	144

Figure 4.66. Train Test Graph of XGBoost + ROS n_estimators 100.....	147
Figure 4.67. Confusion Matrix of XGBoost + ROS n_estimators 100.....	148
Figure 4.68. Classification Report of XGBoost + ROS n_estimators 100	148
Figure 4.69. Train Test Graph of XGBoost + ROS n_estimators 200.....	149
Figure 4.70. Confusion Matrix of XGBoost + ROS n_estimators 200.....	150
Figure 4.71. Classification Report of XGBoost + ROS n_estimators 200	150
Figure 4.72. Train Test Graph of XGBoost + ROS n_estimators 300.....	151
Figure 4.73. Confusion Matrix of XGBoost + ROS n_estimators 300.....	151
Figure 4.74. Classification Report of XGBoost + ROS n_estimators 300	152
Figure 4.75. Train Test Graph of LightGBM + ROS n_estimators 100	153
Figure 4.76. Confusion Matrix of LightGBM + ROS n_estimators 100	153
Figure 4.77. Classification Report of LightGBM + ROS n_estimators 100.....	154
Figure 4.78. Train Test Graph of LightGBM + ROS n_estimators 200	155
Figure 4.79. Confusion Matrix of LightGBM + ROS n_estimators 200	155
Figure 4.80. Classification Report of LightGBM + ROS n_estimators 200.....	156
Figure 4.81. Train Test Graph of LightGBM + ROS n_estimators 300	157
Figure 4.82. Confusion Matrix of LightGBM + ROS n_estimators 300	157
Figure 4.83. Classification Report of LightGBM + ROS n_estimators 300.....	158
Figure 4.84. Train Test Graph of XGBoost + ROS subsample 0.70	160
Figure 4.85. Confusion Matrix of XGBoost + ROS subsample 0.70	161
Figure 4.86. Classification Report of XGBoost + ROS subsample 0.70.....	161
Figure 4.87. Train Test Graph of XGBoost + ROS subsample 0.80	162
Figure 4.88. Confusion Matrix of XGBoost + ROS subsample 0.80	163
Figure 4.89. Classification Report of XGBoost + ROS subsample 0.80.....	163
Figure 4.90. Train Test Graph XGBoost + ROS Subsample 0,9.....	164
Figure 4.91. Confusion Matrix of XGBoost + ROS subsample 0.90	165
Figure 4.92. Classification Report of XGBoost + ROS subsample 0.90.....	165

Figure 4.93. Train Test Graph of LightGBM + ROS subsample 0.70.....	166
Figure 4.94. Confusion Matrix of LightGBM + ROS subsample 0.70.....	167
Figure 4.95. Classification Report of LightGBM + ROS subsample 0.70	167
Figure 4.96. Train Test Graph of LightGBM + ROS subsample 0.80.....	168
Figure 4.97. Confusion Matrix of LightGBM + ROS subsample 0.80.....	169
Figure 4.98. Classification Report of LightGBM + ROS subsample 0.80	169
Figure 4.99. Train Test Graph of LightGBM + ROS subsample 0.90.....	170
Figure 4.100. Confusion Matrix of LightGBM + ROS subsample 0.90.....	170
Figure 4.101. Classification Report of LightGBM + ROS subsample 0.90	171
Figure 4.102. Train Test Graph of XGBoost + ROS colsample_bytree 0.70	174
Figure 4.103. Confusion Matrix of XGBoost + ROS colsample_bytree 0.70	174
Figure 4.104. Classification Report of XGBoost ROS colsample_bytree 0.70...	175
Figure 4.105. Train Test Graph of XGBoost + ROS colsample_bytree 0.80	176
Figure 4.106. Confusion Matrix of XGBoost + ROS colsample_bytree 0.80	176
Figure 4.107. Classification Report of XGBoost ROS colsample_bytree 0.80...	177
Figure 4.108. Train Test Graph of XGBoost + ROS colsample_bytree 0.90	178
Figure 4.109. Confusion Matrix of XGBoost + ROS colsample_bytree 0.90	178
Figure 4.110. Classification Report of XGBoostROS colsample_bytree 0.90....	179
Figure 4.111. Train Test Graph of LightGBM + ROS colsample_bytree 0.70 ...	180
Figure 4.112. Confusion Matrix of LightGBM + ROS colsample_bytree 0.70 ..	180
Figure 4.113. Classification Report LightGBM + ROS colsamplebytree 0.70 ...	181
Figure 4.114. Train Test Graph of LightGBM + ROS colsample_bytree 0.80 ...	182
Figure 4.115. Confusion Matrix of LightGBM + ROS colsample_bytree 0.80 ..	182
Figure 4.116. Classification Report LightGBM + ROS colsample_bytree 0.80 .	183
Figure 4.117. Train Test Graph of LightGBM + ROS colsample_bytree 0.90 ...	184
Figure 4.118. Confusion Matrix of LightGBM + ROS colsample_bytree 0.90 ..	184
Figure 4.119. Classification Report LightGBM + ROS colsample_bytree 0.90 .	185

Figure 4.120. Train Test Graph of XGBoost without ROS	189
Figure 4.121. Confusion Matrix of XGBoost without ROS	189
Figure 4.122. Classification Report of XGBoost without ROS	190
Figure 4.123. Train Test Graph of LightGBM without ROS	191
Figure 4.124. Confusion Matrix of LightGBM without ROS	191
Figure 4.125. Classification Report of LightGBM without ROS	192
Figure 4.126. Login Page Implementation Result	198
Figure 4.127. Register Page Implementation Result.....	199
Figure 4.128. Home Page Implementation Result	200
Figure 4.129. Input Form Page Implementation Result.....	201
Figure 4.131. Model Probability Tab Implementation Result	204
Figure 4.132. Input Summary Tab Implementation Result	205
Figure 4.133. Model Graph Tab Implementation Result	206
Figure 4.134. Children and Adolescents Result Page	207
Figure 4.135. Indonesian Ministry of Health/WHO Table Tab	208

LIST OF TABLES

Table 3.1 Dataset Attribute.....	38
Table 3.2a Preview of Five Rows of the Dhaka Obesity Dataset	39
Table 3.2b Preview of Five Rows of the Dhaka Obesity Dataset (Part B)	39
Table 3.3 Weight Classification Based on BMI for Asia Pacific	41
Table 3.4 Number of Respondents Under 18 Years of Age Removed	42
Table 3.5 Example of Target Relabeling Results Based on Asia Pacific BMI.....	43
Table 3.6 Examples of Column Name Standardization and Variable Codes	44
Table 3.7 Category Value Standardization Examples	45
Table 3.8 Summary of Duplicate Data and Dataset Size	46
Table 3.9 Target Class Distribution Before and After Duplicate Removal	46
Table 3.10 Encoding Methods Used for Each Variable	47
Table 3.11 Example of Binary Variable Mapping into Numbers	47
Table 3.12 Example of Ordered Variable Mapping into Numbers.....	48
Table 3.13 Mapping of Obesity_level Target Classes into Codes 0–4.....	49
Table 3.14 Target Class Distribution Before Balancing.....	50
Table 3.15 General Comparison of XGBoost and LightGBM	52
Table 3.16 Example of Raw Data for XGBoost Model Illustration	56
Table 3.17 Example of Encoded Data for XGBoost Illustration	56
Table 3.18 Gradient and Hessian in the XGBoost Illustration	58
Table 3.19 Split Path and Leaf Weight in XGBoost.....	59
Table 3.20 Updated Score of Each Sample in XGBoost	60
Table 3.21 Illustrative Output of Four Samples After XGBoost Update.....	61
Table 3.22 Illustrative Final Predictions of Four Samples in XGBoost	62
Table 3.23 Example of Raw Data for LightGBM Model Illustration.....	65
Table 3.24 Example of Encoded Data for LightGBM Illustration	65
Table 3.25 Gradient and Hessian in the LightGBM Illustration.....	67

Table 3.26 Split Path and Leaf Weight in LightGBM	69
Table 3.27. Updated Score of Each Sample in LightGBM.....	70
Table 3.28 Illustrative Output of Four Samples After LightGBM Update.....	71
Table 3.29 Illustrative Final Predictions of Four Samples in LightGBM.....	72
Table 3.30 Comparison of Decision Trees in XGBoost and LightGBM.....	73
Table 3.31 Parameter Search Space in the Model Analysis Scheme.....	76
Table 4.1. Main Model Manual Testing Scheme.....	103
Table 4.2. Comparison of XGBoost + ROS Split Ratios.....	110
Table 4.3. Comparison of LightGBM + ROS Split Ratios	118
Table 4.4. Comparison of Learning Rate on XGBoost.....	131
Table 4.5. Comparison of Learning Rate on LightGBM	131
Table 4.6. Comparison of Max Depth on XGBoost	145
Table 4.7. Comparison of Max Depth on LightGBM.....	145
Table 4.8. Comparison of n_estimators on XGBoost.....	158
Table 4.9. Comparison of n_estimators on LightGBM	159
Table 4.10. Comparison of Subsample on XGBoost.....	172
Table 4.11. Comparison of Subsample on LightGBM	172
Table 4.12. Comparison of Colsample_bytree on XGBoost	185
Table 4.13. Comparison of Colsample_bytree on LightGBM.....	186
Table 4.14. Selected Manual Split and Hyperparameters.....	187
Table 4.15. Final Evaluation of XGBoost + ROS and LightGBM + ROS.....	187
Table 4.16. Comparison Results of XGBoost and LightGBM without ROS	193
Table 4.17. Comparison of Models without ROS and with ROS	194
Table 4.18. Performance Difference after Applying ROS.....	195