

**A HYBRID MACHINE LEARNING AND  
OPTIMISATION-BASED MODEL FOR  
PREDICTING THE SUCCESS OF BUSINESS-  
TO-CONSUMER SOFTWARE  
DEVELOPMENT PROJECTS IN INDONESIA**

**RUDI SETIAWAN**

**ASIA e UNIVERSITY**

**2025**

A HYBRID MACHINE LEARNING AND OPTIMISATION-BASED  
MODEL FOR PREDICTING THE SUCCESS OF BUSINESS-TO-  
CONSUMER SOFTWARE DEVELOPMENT PROJECTS IN INDONESIA

RUDI SETIAWAN

A Thesis Submitted to Asia e University in  
Fulfilment of the Requirements for the  
Degree of Doctor of Philosophy

July 2025

## ABSTRACT

Despite the strategic importance of software projects, recent Standish Group findings indicate that only 35% achieve successful completion, with a substantial portion encountering delays or outright failure. Motivated by this challenge, this research aims to advance the predictive capability of software project success by developing an integrated artificial intelligence-based model, contextualized within B2C E-Business software projects in Indonesia. Employing a mixed-methods design, the study first identifies 28 critical success factors through empirical investigation, highlighting elements such as evaluation progress, realistic scheduling, human resource competency, project planning, and risk management. Building on these findings, a predictive framework is constructed by integrating machine learning algorithms with advanced optimization and data handling strategies. Specifically, Support Vector Machines (SVM) enhanced by Aquila Optimization demonstrate superior performance, achieving a predictive accuracy of up to 94.29% on datasets synthesized and balanced through Generative Adversarial Networks (GANs) and hybrid resampling techniques (SMOTE + Tomek Links). To translate predictive insights into actionable guidance, Shapley Additive Explanations (SHAP) are employed to quantify the contribution of each factor, enabling targeted recommendations for mitigating project risks. This study's primary contributions lie in (i) methodologically integrating feature identification, synthetic data augmentation, imbalance handling, hyperparameter optimization, and explainable AI into a cohesive predictive pipeline; (ii) applying this integrated approach within the underexplored context of Indonesian B2C E-Business projects; and (iii) delivering practical decision-support insights that enhance project planning, thereby reducing the likelihood of costly failures. The proposed framework not only elevates predictive accuracy but also reinforces managerial strategies through data-driven prioritization of project factors, offering a robust tool to improve software project success rates in similar contexts.

**Keywords:** Critical success factors, software project, project management, software project prediction, software project success prediction models

## **APPROVAL**

This is to certify that this thesis conforms to acceptable standards of scholarly presentation and is fully adequate, in quality and scope, for the fulfilment of the requirements for the Degree of Doctor of Philosophy.

The student has been supervised by: **Professor Dr Titik Khawa Abdul Rahman**

The thesis has been examined and endorsed by:

**Assistant Professor Ts Dr Nor Hapiza Mohd Ariffin**

Asia e University

Examiner 1

**Associate Professor Ts Dr Saiful Izwan Suliman**

Universiti Teknologi Mara

Examiner 2

This thesis was submitted to Asia e University and is accepted as fulfilment of the requirements for the Degree of Doctor of Philosophy.

*Swa*

**Dr Swa Lee Lee**

Asia e University

Chairperson, Examination Committee

[3 July 2025]

## **DECLARATION**

I hereby declare that the thesis submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy is my own work and that all contributions from any other persons or sources are properly and duly cited. I further declare that the material has not been submitted either in whole or in part, for a degree at this or any other university. In making this declaration, I understand and acknowledge any breaches in this declaration constitute academic misconduct, which may result in my expulsion from the programme and/or exclusion from the award of the degree.

**Name: Rudi Setiawan**

A handwritten signature in black ink, appearing to be 'Rudi Setiawan', written in a cursive style.

**Signature of Student:**

**Date: 3 July 2025**



## **ACKNOWLEDGEMENTS**

With great gratitude and respect, I express my deepest gratitude to Trilogi University for the trust and support I have received through a doctoral study scholarship.

This support is a real form of Trilogi University's commitment to human resource development and a great spirit and motivation for me to continue learning, working, and contributing more widely in the academic field and the wider community.

Specifically, I would also like to express my deepest gratitude and appreciation to my S3 supervisor, Prof. Titik Khawa Binti Abdul Rahman, for the guidance, knowledge, and instructions that are very meaningful in my academic journey. He is not only a guide, but also an inspirational figure who guides me with patience and sincerity.

I would also like to thank the National Research and Innovation Agency (BRIN) Republic of Indonesia for permission to use the MAHAMERU BRIN HPC high-performance computing resource computing service to run models from the results of my research..

## TABLE OF CONTENTS

<b>ABSTRACT</b>	<b>ii</b>
<b>APPROVAL</b>	<b>iii</b>
<b>DECLARATION</b>	<b>iv</b>
<b>ACKNOWLEDGEMENTS</b>	<b>vi</b>
<b>TABLE OF CONTENTS</b>	<b>vii</b>
<b>LIST OF TABLES</b>	<b>x</b>
<b>LIST OF FIGURES</b>	<b>xvii</b>
<b>LIST OF ABBREVIATION</b>	<b>xxvii</b>
<b>CHAPTER 1 INTRODUCTION</b>	<b>1</b>
1.0 Background of Study	1
1.1 Problem Statement	8
1.2 Research Objectives	12
1.3 Research Questions	13
1.4 Conceptual Framework	13
1.5 Justification and Significance of the Study	16
1.6 Operational Definition of Terms	17
1.7 Chapter Summary	20
<b>CHAPTER 2 LITERATURE REVIEW</b>	<b>22</b>
2.0 Introduction	22
2.1 Theoretical Framework	22
2.1.1 Software Project Success Theory	22
2.1.2 Data Augmentation	24
2.1.3 Unbalanced Dataset Handling Techniques	25
2.1.4 Data Standardization	31
2.1.5 Feature Selection	31
2.1.6 Cleaning Data from Outlier	32
2.1.7 Classification Techniques in Machine Learning	33
2.1.8 A Variety of Classification Algorithms	34
2.1.9 Optimization ANN and SVM	40
2.1.10 Optimisation Technique in Machine Learning	42
2.1.11 Multiclass Confusion Matrix	51
2.1.12 Shapley Value Technique for Measuring Contributions of Input Features to Output of a Model (SHAP)	52
2.2 Review of Previous Research	53
2.2.1 Factors for the Success of Software Projects in Indonesia and Other Country	54
2.2.2 Factors Validations Uses Expert Judgement	78
2.2.3 Development of Instrument for Expert Judgement	79
2.2.4 Machine Learning Algorithm for Predicting Software Project	81
2.3 Chapter Summary	89
<b>CHAPTER 3 METHODOLOGY</b>	<b>91</b>
3.0 Introduction	91

3.1	Research Framework	91
3.2	Research Design	97
3.3	Research Method	99
3.3.1	Research Approach	99
3.3.2	Pre-Processing Data	99
3.3.3	Machine Learning Modelling	101
3.3.4	Model Evaluation	108
3.3.5	Implementation and Experimentation	108
3.3.6	Interpretation and Analysis of Results	108
3.3.7	Comparison with Other Models	108
3.4	Data Collection Methods	109
3.5	Expert Validation Process	109
3.6	Identification Significant Influence of Factors and Recommendations for Improvement	110
3.7	Ethical Consideration	111
3.8	Chapter Summary	112
<b>CHAPTER 4 RESULTS</b>		<b>113</b>
4.0	Introduction	113
4.1	Factors Determining the Success of e-Business B2C Software Development Projects in Indonesia	114
4.1.1	Literature Study of Critical Success Factors (CSFs)	115
4.1.2	CSFs Verification Using Expert Judgment	118
4.1.3	Exploratory of Data on Pilot Project Data	123
4.1.4	Feature Selection	149
4.1.5	Exploratory Data Analysis on Dataset Collection Results	156
4.1.6	Handling the Problem of Data Shortages and Dataset Imbalanced	167
4.1.7	Synthetics Dataset Evaluation	180
4.1.8	Handling Imbalanced Dataset	183
4.2	Development of New Classification Techniques in Predicting the Success of e-Business Software Projects in Indonesia	193
4.2.1	Prediction Model for Successful Software Development Project Using ANN	194
4.2.2	Prediction Model for Successful Software Development Project Using SVM	209
4.2.3	Prediction Model for Successful Software Development Project Using Hybrid ANN and Grid Search	222
4.2.4	Prediction Model for Successful Software Development Project Using Hybrid SVM and Grid Search	248
4.2.5	Prediction Model for Successful Software Development Project Using Hybrid ANN and Aquila Optimisation	272
4.2.6	Prediction Model for Successful Software Development Project Using Hybrid SVM and Aquila Optimisation	298
4.2.7	Comparison Results of All Models	323

4.3	Non-Parametric Statistical Test on Accuracy Model	329
4.4	Recommendations for Improving Performance Based on CSFs	331
4.5	Testing on New Data	341
4.6	Chapter Summary	344
<b>CHAPTER 5 DISCUSSION AND CONCLUSION</b>		<b>346</b>
5.0	Introduction	346
5.1	Limitations of Research	347
5.1.1	Dataset Limitations	348
5.1.2	Selection of Machine Learning Algorithms and Optimization Algorithms	348
5.2	Implications of Research	348
5.2.1	Implications in the Academic Field	348
5.2.2	Implications in the Software Industry	349
5.3	Contributions of Research	350
5.3.1	Theoretical Contributions	350
5.3.2	Practical Contributions	351
5.3.3	Contribution to Methodology	351
5.4	Recommendation for Future Research	352
<b>REFERENCES</b>		<b>353</b>
<b>APPENDICES</b>		<b>361</b>
Appendix A		361

## LIST OF TABLES

<b>Table</b>		<b>Page</b>
Table 1.1	Software Project Success Rate	2
Table 1.2	Factors and Sub-Factors for Software Project Success	4
Table 1.3	Percentage of Software Project Success and Failure in Indonesia Based on Project Size	10
Table 2.1	Metrics for a Multiclass Confusion Matrix	52
Table 2.2:	Research on the Factors that Influence Software Project Success	56
Table 2.3	Factors that Influence the Success of Software Projects	58
Table 2.4	Top 5 Most Influential Factors in the Success of Software Projects in Various Cases	65
Table 2.5	Determinant Factors for the Success of Software Projects According to the Literature	68
Table 2.6	Example of a List of Factors and Questions from Garousi's 2019 Research	80
Table 2.7	Previous Research Related to Software Project Failure Detection	84
Table 3.1	Research Design Matrix	93
Table 4.1	Critical Success Factors of Software Project in General	115
Table 4.2	Expert Assessment Results on Success Factors of B2C Software Projects in Indonesia	119
Table 4.3	Mean, Median and Mode in Organizational Factor Group	125
Table 4.4	Variable Statistics x1 – x12	127
Table 4.5	Mean, Median and Mode in Project Factor Groups (Technical and Process)	130
Table 4.6	Statistics of Variables x13-x31	132

Table 4.7	Mean, Median and Mode in Project Manager Factor Groups	136
Table 4.8	Statistics of Variables x32-x36	137
Table 4.9	Mean, Median and Mode in Project Team Factor Groups	139
Table 4.10	Variable Statistics x37-x44	140
Table 4.11	Mean, Median and Mode in External Environmental Factors Groups	143
Table 4.12	Statistics of Variables x45-x 47	144
Table 4.13	Mean, Median and Mode in Customer Factor Groups	146
Table 4.14	Statistics of Variables x48-x51	147
Table 4.15	CSFs Validation Status Based on Pearson Correlation Test	149
Table 4.16	Reliability Test Results	152
Table 4.17	CSFs Ranking Based on Pearson Correlation Test	153
Table 4.18	Statistical Variable Organizational Factors	157
Table 4.19	Statistics Variable Project (Technical and Process) Factors	159
Table 4.20	Variable Statistics Project Manager Factors	162
Table 4.21	Variable Project Team Factors Statistics	164
Table 4.22	Customer Factors Statistics	165
Table 4.23	Comparison of the Number of Original and Synthetic Sample Datasets	167
Table 4.24	Comparison of Original and Synthetic Sample Datasets on Organizational Groups Factors	169
Table 4.25	Statistics of Original and Synthetic Sample Datasets on Project (Technical and Process) Factors	172
Table 4.26	Statistics of Original and Synthetic Sample Datasets on Project Manager Factors	176

Table 4.27	Statistics of Original and Synthetic Sample Datasets on Project Team Factors	177
Table 4.28	Statistics of Original and Synthetic Sample Datasets on Customer Factors	179
Table 4.29	Evaluation Data Augmented Using EMD	180
Table 4.30	Distribution of SMOTE Data Results Before and After Removing Outliers	185
Table 4.31	Distribution of ADASYN Result Data Before and After Removing Outliers	187
Table 4.32	Distribution of Tomek Links Data Results Before and After Removing Outliers	189
Table 4.33	Distribution of SMOTE + Tomek Links Data Results Before and After Removing Outliers	191
Table 4.34	Distribution of ADASYN + Tomek Links Data Results Before and After Removing Outliers	193
Table 4.35	Results of SVM Model on Testing Data from GANs	195
Table 4.36	Results of SVM Model on Testing Data from SMOTE	198
Table 4.37	Results of SVM Model on Testing Data from ADASYN	200
Table 4.38	Results of SVM Model on Testing Data from Tomek Links	203
Table 4.39	Results of SVM Model on Testing Data from SMOTE Tomek Links	205
Table 4.40	Results of SVM Model on Testing Data from ADASYN Tomek Links	208
Table 4.41	Results of SVM Model on Testing Data from GANs	211
Table 4.42	Results of SVM Model on Testing Data from SMOTE	213

Table 4.43	Results of SVM Model on Testing Data from ADASYN	215
Table 4.44	Results of SVM Model on Testing Data from Tomek Links	217
Table 4.45	Results of SVM Model on Testing Data from SMOTE Tomek Links	219
Table 4.46	Results of SVM Model on Testing Data from ADASYN Tomek Links	221
Table 4.47	Parameter Initiation for ANN + Grid Search Model	223
Table 4.48	Results of Hybrid ANN and Grid Search Model on Training Data from GANs	224
Table 4.49	Results of Hybrid ANN and Grid Search Model on Testing Data from GANs	226
Table 4.50	ANN + Grid Search + SMOTE Model Results on Training Data	228
Table 4.51	ANN + Grid Search + SMOTE Model Results on Test Data	230
Table 4.52	ANN + Grid Search + ADASYN Model Results on Training Data	232
Table 4.53	ANN + Grid Search + ADASYN Model Results on Test Data	234
Table 4.54	Results of Hybrid ANN and Grid Search Model on Training Data from Tomek Links	236
Table 4.55	ANN + Grid Search + Tomek Links Model Results on Test Data	238
Table 4.56	ANN + Grid Search and SMOTE + Tomek Links Model Results on Training Data	240
Table 4.57	ANN + Grid Search and SMOTE + Tomek Links Model Results on Test Data	242

Table 4.58	ANN + Grid Search and ADASYN + Tomek Links Model Results on Training Data	244
Table 4.59	ANN + Grid Search and ADASYN + Tomek Links Model Results on Test Data	246
Table 4.60	Solution Search Space for Hyperparameter Tuning in SVM	248
Table 4.61	4.61 SVM + Grid Search + GANs Model Results On Training Data	249
Table 4.62	SVM + Grid Search + GANs Model Results on Test Data	251
Table 4.63	SVM + Grid Search + SMOTE Model Results on Training Data	253
Table 4.64	SVM + Grid Search + SMOTE Model Results on Test Data	255
Table 4.65	SVM + Grid Search + ADASYN Model Results on Training Data	257
Table 4.66	SVM + Grid Search + ADASYN Model Results on Test Data	259
Table 4.67	SVM + Grid Search + Tomek Links Model Results on Training Data	261
Table 4.68	SVM + Grid Search + Tomek Links Model Results on Test Data	263
Table 4.69	SVM + Grid Search and SMOTE + Tomek Links GANs Model Results on Training Data	265
Table 4.70	SVM + Grid Search and SMOTE + Tomek Links Model Results on Test Data	267
Table 4.71	SVM + Grid Search and ADASYN + Tomek Links Model Results on Training Data	269
Table 4.72	SVM + Grid Search and ADASYN + Tomek Links Model Results on Test Data	271

Table 4.73	Solution Search Space for Hyperparameter Tuning in ANN	273
Table 4.74	Test Results with Testing Data on the ANN + AO + GANs Model	276
Table 4.75	Test Results with Testing Data on the ANN + AO + SMOTE Model	280
Table 4.76	Test Results with Testing Data on the ANN + AO + ADASYN Model	284
Table 4.77	Test Results with Testing Data on the ANN + AO + Tomek Links Model	288
Table 4.78	Test Results with Testing Data on the ANN + AO + SMOTE Tomek Links Model	292
Table 4.79	Test Results with Testing Data on the ANN + AO + ADASYN Tomek Links Model	296
Table 4.80	Solution Search Space for Hyperparameter Tuning in SVM	299
Table 4.81	Test Results with Testing Data on the SVM + AO + GANs Model	301
Table 4.82	Test Results with Testing Data on the SVM + AO + SMOTE Model	305
Table 4.83	Test Results with Testing Data on the SVM + AO + ADASYN Model	309
Table 4.84	Test Results with Testing Data on the SVM + AO + Tomek Links Model	313
Table 4.85	Test Results with Testing Data on the SVM + AO + SMOTE Tomek Links Model	318

Table 4.86	Test Results with Testing Data on the SVM + AO + ADASYN Tomek Links Model	322
Table 4.87	Comparison of Software Project Success Prediction Model Performance Results	323
Table 4.88	Recommendations for Improving Performance Based on CSFs	332
Table 4.89	Test Results of Automatic Recommendations Based on SHAP	343

## LIST OF FIGURES

<b>Figure</b>		<b>Page</b>
Figure 1.1	Software Sales Growth in Indonesia	2
Figure 1.2	Conceptual Framework	15
Figure 2.1	Distribution of Oversampling, Undersampling and Class Results from Original Dataset	26
Figure 2.2	SMOTE Pseudocode	27
Figure 2.3	ADASYN Pseudocode	28
Figure 2.4	Tomek Links Pseudocode	29
Figure 2.5	Oversample-BMIs	29
Figure 2.6	SMOTE + Tomek Links Pseudocode	30
Figure 2.7	ADASYN + Tomek Link Pseudocode	30
Figure 2.8	Artificial Neural Network Architecture	35
Figure 2.9	Weight of Each Element and Input and Output of the ANN System	36
Figure 2.10	Hyperplane in SVM	37
Figure 2.11	Shorter Decision Boundary	41
Figure 2.12	Wider Decision Boundary	41
Figure 2.13	Decision Boundary with High Gamma Value	42
Figure 2.14	Decision Boundary with Low Gamma Value	42
Figure 2.15	The Behavior of the Aquila High Soar with the Vertical Stoop	45
Figure 2.16	The Behavior of the Aquila Contour Flight with Short Glide Attack	46
Figure 2.17	The Behavior of the AO In A Spiral Shape	48

Figure 2.18	The Behavior of the Aquila Low Flight with Slow Descent Attack	48
Figure 2.19	The Behavior of the Aquila Walk and Grab Prey	49
Figure 2.20	The Effect of the Quality Function (QF), G <sub>1</sub> ,G <sub>2</sub> on the Behavior of AO	51
Figure 2.21	Multiclass Confusion Matrix	52
Figure 3.1	Research Framework	92
Figure 3.2	Research Design	98
Figure 3.3	Artificial Neural Network Architecture	102
Figure 3.4	Flowchart of hybrid ANN and AO	104
Figure 3.5	Mapping Two-Dimensional Data Space (Left) Into Three-Dimensional Feature Space (Right)	106
Figure 3.6	Flowchart of Hybrid SVM and AO	107
Figure 4.1	Normal Distribution Probability Plot on CSFs Pilot Project Dataset	124
Figure 4.2	Visualization of Data Distribution on Variables x <sub>1</sub> -x <sub>12</sub>	128
Figure 4.3	Visualization of Data Distribution on Variables x <sub>13</sub> -x <sub>31</sub>	134
Figure 4.4	Visualization of Data Distribution on Variables x <sub>32</sub> -x <sub>36</sub>	138
Figure 4.5	Visualization of Data Distribution on Variables x <sub>37</sub> -x <sub>44</sub>	142
Figure 4.6	Visualization of Data Distribution on Variables x <sub>45</sub> -x <sub>47</sub>	145
Figure 4.7	Visualization of Data Distribution on Variables x <sub>48</sub> -x <sub>51</sub>	148
Figure 4.8	Visualization of Data Distribution on Variables Groups Organizational Factors	159
Figure 4.9	Visualization of Data Distribution on Variables Project (Technical and Process) Factors	161

Figure 4.10	Visualization of Data Distribution on Project Manager Factors	163
Figure 4.11	Visualization of Data Distribution on Variables Project Team Factors	165
Figure 4.12	Visualization of Data Distribution on Customer Factors	166
Figure 4.13	Data Distribution in Each Dataset	168
Figure 4.14	Boxplot of Data Distribution on Project Variables (Organization)	171
Figure 4.15	Boxplot of Data Distribution on Project Variables (Technical and Process)	174
Figure 4.16	Boxplot of Data Distribution on Project Variables (Manager)	176
Figure 4.17	Boxplot of Data Distribution on Project Variables (Team)	178
Figure 4.18	Boxplot of Data Distribution on Project Variables (Customer)	179
Figure 4.19	EMD Values for End Variables According to Each Class	182
Figure 4.20	Visualization of Outlier Detection using PCA on SMOTE Result Dataset	184
Figure 4.21	Visualization of Outlier Detection using PCA on ADASYN Result Dataset	186
Figure 4.22	Visualization of Outlier Detection using PCA on Tomek Links Result Dataset	188
Figure 4.23	Visualization of Outlier Detection using PCA on Hybrid SMOTE and Tomek Links Result Dataset	190
Figure 4.24	Visualization of Outlier Detection using PCA on Hybrid ADASYN and Tomek Links Result Dataset	192
Figure 4.25	Accuracy Scores for Each Fold in Cross Validation	195

Figure 4.26	Confusion Matrix on ANN Model from GANs Result Testing Dataset	196
Figure 4.27	Accuracy Scores for Each Fold in Cross Validation	197
Figure 4.28	Confusion Matrix on ANN Model from SMOTE Result Testing Dataset	198
Figure 4.29	Accuracy Scores for Each Fold in Cross Validation	200
Figure 4.30	Confusion Matrix on ANN Model from ADASYN Result Testing Dataset	201
Figure 4.31	Accuracy Scores for Each Fold in Cross Validation	202
Figure 4.32	Confusion Matrix on ANN Model from Tomek Links Result Testing Dataset	203
Figure 4.33	Accuracy Scores for Each Fold in Cross Validation	205
Figure 4.34	Confusion Matrix on ANN Model from SMOTE Tomek Links Result Testing Dataset	206
Figure 4.35	Accuracy Scores for Each Fold in Cross Validation	207
Figure 4.36	Confusion Matrix on ANN Model from ADASYN Tomek Links Result Testing Dataset	208
Figure 4.37	Accuracy Scores for Each Fold in Cross Validation	210
Figure 4.38	Confusion Matrix on SVM Model from GANs Result Testing Dataset	211
Figure 4.39	Accuracy Scores for Each Fold in Cross Validation	212
Figure 4.40	Confusion Matrix on SVM Model from SMOTE Result Testing Dataset	213
Figure 4.41	Accuracy Scores for Each Fold in Cross Validation	214

Figure 4.42	Confusion Matrix on SVM Model from ADASYN Result Testing Dataset	215
Figure 4.43	Accuracy Scores for Each Fold in Cross Validation	216
Figure 4.44	Confusion Matrix on SVM Model from Tomek Links Result Testing Dataset	217
Figure 4.45	Accuracy Scores for Each Fold in Cross Validation	218
Figure 4.46	46 Confusion Matrix on SVM Model from SMOTE Tomek Links Result Testing Dataset	220
Figure 4.47	Accuracy Scores for Each Fold in Cross Validation	221
Figure 4.48	Confusion Matrix on SVM Model from ADASYN Tomek Links Result Testing Dataset	222
Figure 4.49	Accuracy for All Folds on Hybrid ANN and Grid Search Model on Training Data from GANs	225
Figure 4.50	Mean Test Accuracy for Each Match on Hybrid ANN and Grid Search Model on Training Data from GANs	226
Figure 4.51	Confusion Matrix on Hybrid Model of ANN and Grid Search from GANs Result Testing Dataset	227
Figure 4.52	Accuracy for All Folds on ANN + Grid Search + SMOTE Model	229
Figure 4.53	Mean Test Accuracy for Each Match on ANN + Grid Search + SMOTE Model	230
Figure 4.54	Confusion Matrix ANN + Grid Search + SMOTE Model	231
Figure 4.55	Accuracy for All Folds on ANN + Grid Search + ADASYN Model	233

Figure 4.56	Mean Test Accuracy for Each Match on ANN + Grid Search + ADASYN Model	234
Figure 4.57	Confusion Matrix ANN + Grid Search + ADASYN Model	235
Figure 4.58	Accuracy for All Folds on ANN + Grid Search + Tomek Links Model	237
Figure 4.59	Mean Test Accuracy for Each Match on ANN + Grid Search + Tomek Links Model	238
Figure 4.60	Confusion Matrix ANN + Grid Search + Tomek Links Model	239
Figure 4.61	Accuracy for All Folds on ANN + Grid Search and SMOTE + Tomek Links Model	241
Figure 4.62	Mean Test Accuracy for Each Match on ANN + Grid Search and SMOTE + Tomek Links Model	242
Figure 4.63	Confusion Matrix ANN + Grid Search and SMOTE+ Tomek Links Model	243
Figure 4.64	Accuracy for All Folds on ANN + Grid Search and ADASYN + Tomek Links Model	245
Figure 4.65	Mean Test Accuracy for Each Match on ANN + Grid Search and ADASYN + Tomek Links Model	246
Figure 4.66	Confusion Matrix ANN + Grid Search and ADASYN + Tomek Links Model	247
Figure 4.67	Accuracy for All Folds on SVM + Grid Search + GANs Model	250
Figure 4.68	Mean Test Accuracy for Each Match on SVM + Grid Search + GANs Model	251
Figure 4.69	Confusion Matrix SVM + Grid Search + GANs Model	252

Figure 4.70	Accuracy for All Folds on SVM + Grid Search + SMOTE Model	254
Figure 4.71	Mean Test Accuracy for Each Match on SVM + Grid Search + SMOTE Model	255
Figure 4.72	Confusion Matrix SVM + Grid Search + SMOTE Model	256
Figure 4.73	Accuracy for All Folds on SVM + Grid Search + ADASYN Model	258
Figure 4.74	Mean Test Accuracy for Each Match on SVM + Grid Search + ADASYN Model	259
Figure 4.75	Confusion Matrix SVM + Grid Search + ADASYN Model	260
Figure 4.76	Accuracy for All Folds on SVM + Grid Search + Tomek Links Model	262
Figure 4.77	Mean Test Accuracy for Each Match on SVM + Tomek Links Model	263
Figure 4.78	Confusion Matrix SVM + Grid Search + Tomek Links Model	264
Figure 4.79	Accuracy for All Folds on SVM + Grid Search and SMOTE + Tomek Links Model	266
Figure 4.80	Mean Test Accuracy for Each Match on SVM + Grid Search and SMOTE + Tomek Links Model	267
Figure 4.81	Confusion Matrix SVM + Grid Search and SMOTE + Tomek Links Model	268
Figure 4.82	Accuracy for All Folds on SVM + Grid Search and ADASYN + Tomek Links Model	270
Figure 4.83	Mean Test Accuracy for Each Match on SVM + Grid Search and ADASYN + Tomek Links Model	271