

**PREDICTING CUSTOMER CHURN IN INDONESIAN ISPs: AN MLP AND
MARKETING INTELLIGENCE APPROACH - A CASE STUDY ON PT XYZ**

By

GEMA PERSADA ARIHTA
[22252010]

MASTER'S DEGREE
in

MECHANICAL ENGINEERING – ENGINEERING MANAGEMENT
FACULTY OF ENGINEERING AND INFORMATION OF TECHNOLOGY



SWISS GERMAN UNIVERSITY
The Prominence Tower
Jalan Jalur Sutera Barat No. 15, Alam Sutera
Tangerang, Banten 15143 - Indonesia

Revision after the Thesis Defense on 22nd Januari 2024

STATEMENT BY THE AUTHOR

I hereby declare that this submission is my own work and to the best of my knowledge, it contains no material previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any other degree or diploma at any educational institution, except where due acknowledgement is made in the thesis.

Gema Persada Arihta

Student

Date

Approved by:

Assoc. Prof. Dr. Tanika D. Sofianti, S.T., M.T

Thesis Advisor

Date

Dr. Ir. Win Sukardi M. Eng. M.M. MBA M. Hum

Thesis Co-Advisor

Date

Dr. Ir. Maulahikmah Galinium, S.Kom., M.Sc

Dean Faculty Engineering and Information Technology

Date

ABSTRACT

PREDICTING CUSTOMER CHURN IN INDONESIAN ISPs: AN MLP AND MARKETING INTELLIGENCE APPROACH - A CASE STUDY ON PT XYZ

By

Gema Persada Arihta

Assoc. Prof. Dr. Tanika D. Sofianti, S.T., M.T, Advisor

Dr. Ir. Win Sukardi M. Eng. M.M. MBA M. Hum, Co-Advisor

SWISS GERMAN UNIVERSITY

To forecast customer churn in the Indonesian Internet service provider (ISP) industry, this research proposes a unique combination of Multilayer Perceptron (MLP) neural networks with marketing intelligence and strategy. This research examines various aspects of customer retention issues using a comprehensive dataset from a top Indonesian ISP that includes user demographics, subscription information, geographic information, and customer interaction.

The research focuses on using MLP, a sophisticated artificial neural network, to predict churn potential, providing a deeper understanding of churn dynamics by examining key factors influencing customer preferences. With the expected results for above 90% accuracy rate in predicting customer churn, this research will provide a valuable marketing insights for the company. This will enable PT XYZ to tailor customer engagement and loyalty campaigns by combining machine learning (MLP) with targeted marketing approaches. This combination can lead to reduced churn rates, improved customer satisfaction, and a long-term competitive advantage.

The research highlights the importance of advanced data analytics in enhancing telecom marketing techniques. It emphasizes the need for machine learning to understand consumer behaviour and develop proactive marketing strategies to maintain customer loyalty in the competitive Indonesian ISP industry. The study aims to provide ISPs with valuable insights into customer behaviour, enabling them to develop effective retention strategies and maintain a competitive edge in a constantly evolving industry.

Keywords: Customer Churn, Telecommunications Industry, Marketing Intelligence, Predictive Modelling, Case Study, PT XYZ, Indonesia.

© Copyright 2024
by Gema Persada Arihta
All rights reserved

DEDICATION

I dedicate this work to my family, all my friends, and for Indonesia

Also, to Management PT Supra Primatama Nusantara

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my thesis advisor, Assoc. Prof. Dr. Tanika D. Sofianti, S.T., M.T., and my thesis co-advisor Dr. Ir. Win Sukardi M. Eng. M.M. MBA M. Hum, for their invaluable guidance, expertise, and unwavering patience. Your mentorship has been instrumental in shaping this research and my academic growth.

Also, to my family, thank you for the passion and prayers that always accompany the steps to complete the study master program. Thanks to my friends and loved ones who have been a source of inspiration, motivation, and understanding. Your belief in me has kept me going during challenging times.

Lastly, thanks to the management of PT Supra Primatama Nusantara who have given me support to take a master's degree at Swiss German University.

TABLE OF CONTENTS

	Page
STATEMENT BY THE AUTHOR	2
ABSTRACT	3
DEDICATION	6
ACKNOWLEDGEMENTS	7
TABLE OF CONTENTS	8
LIST OF FIGURES	10
LIST OF TABLES	12
CHAPTER 1 - INTRODUCTION	13
1.1 Background	13
1.2 Research Problems.....	18
1.3 Research Objectives.....	20
1.4 Significance of Study.....	20
1.5 Research Questions	21
1.6 Expected Outcome	21
CHAPTER 2 - LITERATURE REVIEW	22
2.1 Introduction to Customer Churn	22
2.2 Churn Prediction Models	24
2.3 Machine Learning in Churn Prediction.....	25
2.4 Marketing Strategies to Mitigate Churn	27
2.5 Research Comparison	28
CHAPTER 3 – RESEARCH METHOD	31
3.1 Data Description	32
3.2 Data Preparation and Preprocessing	33
3.3 MLP Neural Network Design	33
3.4 Exploratory Data Analysis (EDA)	34
3.5 Model Evaluation.....	35
3.6 Integration of MLP and EDA Findings	36
CHAPTER 4 – RESULTS AND DISCUSSION	37
4.1 Data Overview	37
4.2 Data Preparation and Preprocessing	39

4.3 Exploratory Data Analysis (EDA)	50
4.4 Model Design	79
4.5 Model Evaluation	90
4.6 Integration of MLP and EDA	93
4.7 Findings	96
4.8 Deployment	100
CHAPTER 5 – CONCLUSIONS AND RECCOMENDATIONS	101
5.1 Conclusions	101
5.2 Recommendations	102
REFERENCES	104
CURRICULUM VITAE	109

LIST OF FIGURES

Figures	Page
Figure 1.1. PT XYZ Termination Rate.	13
Figure 1.2. Customer New Revenue and Termination Revenue Loss at PT XYZ	19
Figure 1.3. New and Termination Customers at PT XYZ	19
Figure 2.1 Multiple-Layer Perceptron Architectures (Da Silva et al., 2016).....	26
Figure 3.1 Research Methodology	31
Figure 4.1 Data Duplication in Dataset.....	40
Figure 4.2 Data Duplication Result and New Feature	40
Figure 4.3 Missing Values on Dataset	41
Figure 4.4 Data Type Conversion	43
Figure 4.5 Package Price Investigation.....	47
Figure 4.6 <i>call_status</i> Feature Values.....	49
Figure 4.7 <i>category</i> Feature Values	49
Figure 4.8 <i>prepaidtype</i> Values	50
Figure 4.9 <i>cust_seg</i> Values Result	50
Figure 4.10 Customer Churn by Gender.....	51
Figure 4.11 Churn Distribution by Age Group	54
Figure 4.12 Boxplot of Package Price and Churn.....	55
Figure 4.13 Top 20 Popular Package Name	58
Figure 4.14 Top 20 Package Name by Churn Rate	58
Figure 4.15 Distribution of Subscription Tenure	60
Figure 4.16 Top 20 Profit Centers Name	61
Figure 4.17 Distribution of Total Tickets for Churned and Non-Churned Customers	63
Figure 4.18 Churn Rate by Call Status	64
Figure 4.19 Distribution of <i>seg</i> Reason for Call Failed Status	65
Figure 4.20 Churn by Category.....	66
Figure 4.21 Distribution of <i>seg</i> Reasons Within "customer" Category	67
Figure 4.22 Number of Customers by Province	68
Figure 4.23 Churned Customers by Province	69
Figure 4.24 Churn Rate by Province.....	70

Figure 4.25 Customer Churn by Territory	71
Figure 4.26 Percentage Customer Churn by Territory	71
Figure 4.27 Customer Churn by Area Type (Urban/Rural)	72
Figure 4.28 Customer Churn Percentage by Area Type (Urban/Rural)	73
Figure 4.29 Customer Churn by Technology	74
Figure 4.30 Customer Churn by Customer Segment	76
Figure 4.31 Customer Churn Rate Percentage by Customer Segment	76
Figure 4.32 Churn Distribution	78
Figure 4.33 Churn Rate Percentage	78
Figure 4.34 Correlation Matrix for Numerical Features and Target	80
Figure 4.35 Class Distribution Before Sampling (SMOTE)	83
Figure 4.36 Class Distribution After SMOTE	84
Figure 4.38 Train and Validation Model Test with and Without Initial Bias	86
Figure 4.39 MLP Model Architecture	87
Figure 4.40 SMOTE Confusion Matrix Result	91

LIST OF TABLES

Table	Page
Table 1.1 Machine Learning Techniques Comparison (Vafeiadis et al., 2015; Zhang et al., 2023)	15
Table 1.2 The Use of Machine Learning in Marketing Intelligence (Jahan & Farah Sanam, 2022; Shaker Reddy et al., 2023; Sonatha, 2013; T.Ş Gürsoy, 2010)	17
Table 2.1 Factors Impacting Customer Churn (Salma & Aprianingsih Ph.D, 2021b; Triyafebrianda & Windasari, 2022)	23
Table 2.2 Churn Prediction Model Comparison (Bogaert & Delaere, 2023; Geiler et al., 2022; Sana et al., 2022)	25
Table 2.3 Research Comparison	30
Table 3.1 PT XYZ Dataset Information	32
Table 4.1 PT XYZ Details Dataset Information	38
Table 4.2 Standardize Feature and Feature Selection Result	39
Table 4.3 Filling <i>sub1</i> Missing Value	42
Table 4.4 Filling <i>seg</i> Missing Value	42
Table 4.5 Numerical Data	43
Table 4.6 Categorical Data	44
Table 4.7 Customer with Age Under 17	45
Table 4.8 Customer with Age Over 80	45
Table 4.9 Subscription Tenure	47
Table 4.10 Package Price Investigation Result	48
Table 4.11 Call Failed Value	49
Table 4.12 Feature Dataset	51
Table 4.13 Feature Selection for Modeling	79
Table 4.14 Numeric ANOVA Test	81
Table 4.15 Chi-Square Categorical Result	82
Table 4.16 Model Test with and Without Initial Bias	85
Table 4.17 Model Performance Results	89
Table 4.18 SMOTE Performance Metrics	90