

LIST OF CONTENTS

ABSTRACT.....	i
APPROVAL PAGE	ii
ORIGINALITY STATEMENT SHEET.....	iii
PREFACE	iv
LIST OF CONTENTS	v
LIST OF FIGURES	ix
LIST OF TABLES	x
GLOSSARY.....	xi
LIST OF ABBREVIATIONS	xii
LIST OF APPENDICES	xiii
Chapter I INTRODUCTION	1
I.1 Background.....	1
I.2 Problem Formulation	4
I.3 Research Objective	4
I.4 Research Benefits	5
I.5 Research Limitations and Assumptions.....	6
I.6 Structure of Report.....	6
Chapter II LITERATURE REVIEW.....	8
II.1 Mental Health and Mental Health Problems.....	8
II.1.1 Depression	9
II.1.2 Anxiety	9
II.2 Natural Language Processing	10
II.3 CRISP-DM.....	10
II.4 Text Pre-processing	12
II.4.1 Data Cleaning	12

II.4.2	Case Folding.....	13
II.5	Tokenization	13
II.6	Data Balancing.....	14
II.7	Multi-Label Classification	15
II.8	Transformer Model	16
II.8.1	Encoder and Decoder	17
II.8.2	Attention.....	17
II.8.3	Feed-Forward Network	19
II.8.4	Linear Output Layer in Classification	20
II.8.5	Sigmoid Activation in Multi-label Classification	20
II.9	Fine-tuned Pre-trained Models	21
II.9.1	Bidirectional Encoder Representations from Transformers (BERT) 21	
II.9.1.1	IndoBERT.....	22
II.9.1.2	Input Embedding	23
II.10	Binary Cross-Entropy Loss for Multi-Label Classification.....	24
II.11	Model Performance Evaluation Method.....	25
II.12	Methodology Comparison	27
II.13	State of The Art.....	28
Chapter III	PROBLEM-SOLVING METHODOLOGY	38
III.1	Research Methodology	38
III.2	Systematic Approach of Research	39
Chapter IV	PROBLEM RESOLUTION.....	43
IV.1	Business Understanding.....	43
IV.2	Data Understanding	44
IV.2.1	Login Session Saving	45

IV.2.2	Facebook Data Collection Process	45
IV.2.3	TikTok Data Collection Process.....	47
IV.2.4	Data Exploration and Quality Assessment.....	48
IV.3	Data Preparation	49
IV.3.1	Text Preprocessing	49
IV.3.2	Data Labelling	52
IV.3.3	Data Splitting.....	54
IV.3.4	Data Balancing	55
IV.3.5	Tokenization	56
IV.4	Modelling.....	58
IV.4.1	Model Configuration and Hyperparameter Setup	58
IV.4.2	Hyperparameter Tuning via Grid Search	60
IV.4.2.1	Grid Search Methodology	61
IV.4.2.2	Training Methodology	61
IV.4.2.3	Training with Original Dataset.....	63
IV.4.2.4	Training with Random Oversampling (ROS) Dataset	65
IV.4.2.5	Training with Random Undersampling (RUS) Dataset	67
IV.4.3	Model Validation and Selection	69
IV.4.3.1	Validation Method	69
IV.4.3.2	Performance Comparison of Models	70
IV.4.3.3	Convergence Analysis.....	71
IV.4.3.4	Best Model Selection	72
IV.4.3.5	Mathematical Model Example	72
Chapter V	VALIDATION, RESULT ANALYSIS, AND IMPLICATIONS	75
V.1	Evaluation	75
V.2	Deployment.....	77

V.3	Result Evaluation.....	79
V.4	Final Project Result Implications.....	82
Chapter VI	CONCLUSION AND RECOMMENDATION.....	83
VI.1	Conclusion	83
VI.2	Recommendation	84
APPENDIX		86
REFERENCES		94