

---

# CONTENTS

<b>APPROVAL</b>	<b>ii</b>
<b>SELF DECLARATION AGAINST PLAGIARISM</b>	<b>iii</b>
<b>ABSTRACT</b>	<b>iv</b>
<b>ABSTRAK</b>	<b>v</b>
<b>DEDICATION</b>	<b>vi</b>
<b>ACKNOWLEDGMENTS</b>	<b>vii</b>
<b>PREFACE</b>	<b>viii</b>
<b>CONTENTS</b>	<b>ix</b>
<b>LIST OF TABLES</b>	<b>xi</b>
<b>LIST OF FIGURES</b>	<b>xiii</b>
<b>LIST OF NOTATIONS</b>	<b>xiv</b>
<b>1 INTRODUCTION</b>	<b>1</b>
1.1 Rationale . . . . .	1
1.2 Statement of the Problem . . . . .	2
1.3 Objective and Hypotheses . . . . .	3
1.4 Assumption . . . . .	4
1.5 Scope and Delimitation . . . . .	4
1.6 Significance of the Study . . . . .	5
<b>2 REVIEW OF LITERATURE AND STUDIES</b>	<b>6</b>
2.1 Related Studies . . . . .	6
2.2 Related Literatures . . . . .	9
2.2.1 Abusive Language . . . . .	9
2.2.2 Hate Speech . . . . .	9
2.2.3 Data Balancing . . . . .	9
2.2.4 Transformer . . . . .	11
2.2.5 Fine-tuned Pre-trained Models . . . . .	14
2.2.6 Bidirectional Encoder Representations from Transformers (BERT) .	14
2.2.7 Robustly Optimizer Bidirectional BERT Pre-training Approach (RoBERTa)	17
2.2.8 K-Fold Cross Validation . . . . .	18

2.2.9	Model Performance Evaluation Metrics . . . . .	19
<b>3</b>	<b>RESEARCH METHODOLOGY</b>	<b>20</b>
3.1	Research Design . . . . .	20
3.1.1	Dataset . . . . .	21
3.1.2	Text Preprocessing . . . . .	21
3.1.3	Data Splitting . . . . .	22
3.2	Experiment Scenario . . . . .	22
3.3	Tools for Data Analysis . . . . .	24
3.3.1	Programming Language . . . . .	24
3.3.2	Library dan Framework . . . . .	24
3.3.3	Tools . . . . .	25
<b>4</b>	<b>PRESENTATION, ANALYSIS AND INTERPRETATION OF DATA</b>	<b>26</b>
4.1	Presentation of Data . . . . .	27
4.1.1	Training and Hyperparameter Tuning . . . . .	27
4.1.2	Comparison the Model without Data Balancing . . . . .	27
4.1.3	Comparison the Model with Data Balancing . . . . .	29
4.1.4	Confusion Matrix for Analysis and Evaluation . . . . .	31
4.1.5	Comparison with Other Studies . . . . .	32
4.2	Analysis of the Data . . . . .	34
4.2.1	Training and Hyperparameter Tuning . . . . .	34
4.2.2	Comparison the Model without Data Balancing . . . . .	34
4.2.3	Comparison the Model with Data Balancing . . . . .	35
4.2.4	Confusion Matrix for Analysis and Evaluation . . . . .	36
4.2.5	Comparison with Other Studies . . . . .	37
4.3	Summary of Findings . . . . .	38
<b>5</b>	<b>CONCLUSION AND RECOMMENDATIONS</b>	<b>42</b>
5.1	Conclusions . . . . .	42
5.2	Recommendations . . . . .	42
	<b>BIBLIOGRAPHY</b>	<b>44</b>
	<b>Appendices</b>	<b>48</b>
<b>A</b>	<b>MISCELLANEOUS</b>	<b>50</b>
A.1	Comparison the Model without Data Balancing . . . . .	50
A.2	Comparison the Model with Data Balancing . . . . .	54
<b>B</b>	<b>Curriculum Vitae # Example</b>	<b>67</b>