

## CONTENTS

<b>HALAMAN PENGESAHAN</b>	<b>iii</b>
<b>HALAMAN PENGESAHAN</b>	<b>iv</b>
<b>DECLARATION PAGE</b>	<b>v</b>
<b>HALAMAN PERNYATAAN</b>	<b>vi</b>
<b>TRIBUTE</b>	<b>vii</b>
<b>MOTTO PAGE</b>	<b>viii</b>
<b>PREFACE</b>	<b>ix</b>
<b>CONTENTS</b>	<b>xi</b>
<b>LIST OF SYMBOLS</b>	<b>xiii</b>
<b>ABSTRACT</b>	<b>xvi</b>
<b>INTISARI</b>	<b>xvii</b>
<b>I Introduction</b>	<b>1</b>
1.1. Background	1
1.2. Problem Formulation	1
1.3. Research Objectives and Benefits	2
1.4. Literature Review	2
1.5. The Method of Research	3
1.6. Systematic Approach of Writing	4
<b>II Fundamental Theory</b>	<b>5</b>
2.1. Loss Function	5
2.2. Stochastic Gradient Descent with Momentum	6
2.3. Neural Networks	8
2.3.1. Definition of Neural Networks	8
2.3.2. Architecture of Neural Networks	10
2.3.3. Example of neural networks	12
2.4. Word Embedding	20
2.5. Recurrent Neural Networks	21
2.5.1. Architecture of Recurrent Neural Network	22
2.5.2. Training a Recurrent Neural Network	23
2.5.3. Types of RNNs	24
2.5.4. Bidirectional Recurrent Neural Networks (Bi-RNN)	26
2.6. Softmax function	26
2.7. Matrix Calculus	27

2.7.1. Matrix Differentiation . . . . .	27
2.7.2. Hadamard Product . . . . .	30
2.8. Probability Theory . . . . .	31
2.8.1. Conditional Probability . . . . .	31
2.8.2. Independence . . . . .	32
<b>III Smart Reply Application Structure . . . . .</b>	<b>33</b>
3.1. Long Short-Term Memory (LSTM) . . . . .	33
3.1.1. Architecture of Long Short-Term Memory . . . . .	34
3.1.2. Backpropagation of weights and biases of LSTM . . . . .	40
3.1.3. Bidirectional LSTM . . . . .	60
3.2. Sequence-to-Sequence Architectures . . . . .	61
3.3. Beam Search . . . . .	63
<b>IV RESULT . . . . .</b>	<b>68</b>
4.1. Result 1 . . . . .	71
4.2. Result 2 . . . . .	77
4.3. Result 3 . . . . .	82
4.4. Result 4 . . . . .	87
<b>V SUMMARY . . . . .</b>	<b>93</b>
5.1. Conclusion . . . . .	93
5.2. Recommendation . . . . .	94
<b>REFERENCES . . . . .</b>	<b>95</b>