

## TABLE OF CONTENTS

<b>UNDERGRADUATE THESIS.....</b>	<b>i</b>
<b>APPROVAL PAGE.....</b>	<b>ii</b>
<b>PLAGIARISM STATEMENT.....</b>	<b>iii</b>
<b>FOREWORD.....</b>	<b>iv</b>
<b>TABLE OF CONTENTS.....</b>	<b>v</b>
<b>LIST OF TABLES.....</b>	<b>viii</b>
<b>LIST OF FIGURES.....</b>	<b>ix</b>
<b>ABSTRACT.....</b>	<b>xi</b>
<b>CHAPTER I INTRODUCTION.....</b>	<b>1</b>
1.1 Research Background.....	1
1.2 Research Problem.....	3
1.3 Research Scope.....	4
1.4 Research Objective.....	4
1.5 Research Benefit.....	5
<b>CHAPTER II LITERATURE REVIEW.....</b>	<b>6</b>
<b>CHAPTER III THEORETICAL BASIS.....</b>	<b>15</b>
3.1 Natural Language Processing.....	15
3.2 Information Retrieval.....	16
3.3 FL3XX Dataset.....	17
3.4 Autocomplete & Autocorrect.....	18
3.4.1 Edit Distance.....	20
3.5 Trie Algorithm.....	21
3.5.1 Radix Trie Algorithm.....	22
3.5.2 Hashmap-Based Trie Algorithm.....	25
3.6 Performance Evaluation.....	30
3.6.1 Space and Time Complexity.....	31
3.6.2 Prefix and Suffix Score.....	31
3.6.3 Similarity Score.....	32
3.6.4 Overall Score of Dictionary Terms.....	33
3.6.5 Mean Reciprocal Rank.....	33
<b>CHAPTER IV RESEARCH METHODOLOGY.....</b>	<b>35</b>
4.1 Research Description.....	35
4.2 Data Collection.....	37
4.3 Data Preprocessing.....	38
4.4 Model Development.....	40
4.4.1 Radix Tries.....	40

4.4.2 Hashmap-based Tries.....	48
4.5 Autocomplete & Autocorrect.....	52
4.6 Performance Evaluation.....	56
4.6.1 Prefix Score.....	57
4.6.2 Suffix Score.....	57
4.6.3 Similarity Score.....	58
4.6.4 Overall Score.....	59
4.6.5 Scenario 1: Real-World Data.....	60
4.6.6 Scenario 2: Controlled Testing.....	60
4.7 Model Evaluation.....	61
<b>CHAPTER V IMPLEMENTATION.....</b>	<b>62</b>
5.1 Hardware and Software Specification.....	62
5.2 Data Collection.....	62
5.3 Data Preparation.....	63
5.3.1 Train Data.....	63
5.3.2 Test Data.....	63
5.4 Model Development.....	65
5.4.1 Radix Tries.....	65
5.4.2 Hashmap-based Tries.....	68
5.5 Autocomplete & Autocorrect.....	70
5.6 Performance Evaluation.....	73
5.6.1 Prefix Score.....	73
5.6.2 Suffix Score.....	73
5.6.3 Similarity Score.....	74
5.7 Model Evaluation.....	75
5.7.1 Overall Similarity Score.....	75
5.7.2 Time Performance.....	75
5.7.3 Space Complexity.....	76
5.7.4 MRR.....	76
<b>CHAPTER VI RESULTS AND DISCUSSION.....</b>	<b>78</b>
6.1 Generated Dataset Results.....	78
6.2 Performance Evaluation.....	79
6.2.1 Scenario 1: Real-World Data Results.....	79
6.2.2 Scenario 2: Controlled Testing.....	81
6.3 Model Evaluation.....	83
6.4 User Interface Preview.....	90
<b>CHAPTER VII CONCLUSION AND SUGGESTION.....</b>	<b>92</b>
7.1 Conclusion.....	92

7.2 Future Works.....	93
<b>References.....</b>	<b>94</b>