

LIST OF CONTENTS

ABSTRACT.....	i
APPROVAL PAGE	ii
ORIGINALITY STATEMENT SHEET	iii
PREFACE	iv
LIST OF CONTENTS	v
LIST OF FIGURES	viii
LIST OF TABLES	x
Chapter I INTRODUCTION	1
I.1 Background	1
I.2 Problem Formulation.....	4
I.3 Research Objective.....	4
I.4 Research Benefits	4
I.5 Research Limitation	5
I.6 Report Structure	6
Chapter II Literature review	8
II.1 Mental Health.....	8
II.2 Natural Language Processing (NLP).....	9
II.3 CRISP-DM	10
II.4 Machine Learning	12
II.5 Text Preprocessing	12
II.6 Word2Vec	16

II.7	K-Means	16
II.7.1	Euclidean Distance	18
II.7.2	Cluster Assignment	19
II.7.3	Centroid	19
II.7.4	Elbow Method	20
II.8	Evaluation Method	21
II.8.1	Silhouette Index	21
II.8.2	Calinski-Harabasz Index	21
II.9	Methodology comparison	22
II.10	State of the Art	23
Chapter III	Problem solving methodology	29
III.1	Research Methodology	29
III.2	Systematic Problem Solving	30
Chapter IV	Problem resolution	33
IV.1	Business Understanding	33
IV.2	Data Understanding	33
IV.3	Data Preparation	35
IV.3.1	Text Preprocessing	36
IV.4	Modeling	40
IV.4.1	Word2Vec	41
IV.4.2	K - Means	44
Chapter V	Validation, Result Analysis, and Implication	52
V.1	Evaluation and Validation	52

V.1.1	Evaluation.....	52
V.1.2	Topic Cluster Analysis	54
V.1.3	Cluster 0 Topic Analysis	55
V.1.4	Cluster 1 Topic Analysis	57
V.1.5	Cluster 2 Topic Analysis	58
V.1.6	Cluster 3 Topic Analysis	59
V.1.7	Validation	62
V.2	Deployment	64
V.3	Result Evaluation	66
V.4	Implication	67
Chapter VI	Conclusion and suggestions	69
VI.1	Conclusion.....	69
VI.2	Suggestion	70
References		72