

TABLE OF CONTENT

COVER	i
APPROVAL PAGE	ii
STATEMENT	iii
FOREWORD	iv
TABLE OF CONTENT	vi
LIST OF FIGURES.....	ix
LIST OF TABLES	xi
ABSTRACT.....	xii
INTISARI.....	xiii
CHAPTER I INTRODUCTION	1
1.1 Research Background.....	1
1.2 Research Problem.....	3
1.3 Research Scope.....	3
1.4 Research Objectives	3
1.5 Research Benefits	4
1.6 Research Methodology	4
1.7 Thesis Organization.....	5
CHAPTER II LITERATURE REVIEW.....	7
CHAPTER III THEORETICAL BASIS.....	13
3.1 Recommender System.....	13
3.2 Collaborative Filtering.....	13
3.2.1 User-based Collaborative Filtering	14
3.2.2 Memory- based Collaborative Filtering Issues	14
3.3 Similarity Measures.....	15
3.3.1 Pearson Correlation.....	15
3.3.2 Proximity-Impact-Popularity	16
3.4 Fuzzy C-Means Clustering	19
3.4.1 Procedures of Fuzzy C-Means Clustering.....	19
3.4.2 Fuzzy C-Means Validity Index	22
3.5 Evaluation.....	22
3.5.1 K-Fold Cross Validation	23
3.5.2 Mean Absolute Error.....	23
3.5.3 Coverage	23
3.5.4 Throughput.....	24
CHAPTER IV ANALYSIS AND DESIGN	25
4.1 System Analysis	25
4.1.1 System Description	25
4.1.2 Data Description.....	29
4.2 System Design	31
4.2.1 Data Split.....	31
4.2.2 Data Transformation	32
4.2.3 Fuzzy C-Means Clustering.....	33
4.2.4 Neighbor Finding	40
4.2.5 Similarity Measures	44

4.2.6	Weighted Average for Rating Prediction.....	47
4.2.7	Performance Evaluation.....	47
CHAPTER V	IMPLEMENTATION.....	51
5.1	Implementation Environment.....	51
5.2	Implementation of Data Split.....	51
5.3	Implementation of Fuzzy C-Means Clustering.....	52
5.3.1	Data input and Parameter Setting.....	52
5.3.2	Membership Partition Matrix Initialization.....	53
5.3.3	Cluster Center Calculation.....	53
5.3.4	Euclidean Distance Calculation.....	54
5.3.5	Objective Function Calculation.....	55
5.3.6	Update Membership Partition Matrix Calculation.....	55
5.3.7	Overall Fuzzy C-Means Function Implementation.....	56
5.3.8	Cluster Validity Index.....	57
5.4	Implementation of Data Reading.....	57
5.5	Neighbor Finding.....	58
5.5.1	Neighbor Finding on User-based CF.....	58
5.5.2	Neighbor Finding on User-based CF with Fuzzy C-Means Best Cluster Defuzzification.....	58
5.5.3	Neighbor Finding on User-based CF with Fuzzy C-Means All Cluster Defuzzification.....	59
5.6	Implementation of Similarity Measures.....	60
5.6.1	Pearson Correlation.....	60
5.6.2	Proximity-Impact-Popularity.....	61
5.7	Weighted Average For Rating Prediction.....	64
5.7.1	Rating Prediction in User-based CF.....	64
5.7.2	Rating Prediction in User-based CF with Fuzzy C-Means Best Cluster Defuzzification.....	66
5.7.3	Rating Prediction in User-based CF with Fuzzy C-Means All Cluster Defuzzification.....	67
5.8	Perfarmce Evaluation.....	68
CHAPTER VI	RESULT AND DISCUSSION.....	73
6.1	The Parameter Selection on Fuzzy C-Means.....	73
6.2	The Implementation Result.....	74
6.2.1	The Implementation Result of User-based CF.....	74
6.2.2	The Implementation Result of User-based CF with Fuzzy C-Means Best Cluster Defuzzification.....	77
6.2.3	The Implementation Result of User-based CF with Fuzzy C-Means All Cluster Defuzzification.....	79
6.3	The Performance Evaluation Result on Sparsity, Scalability, and Cold-Start Problem.....	81
6.3.1	The Parameters Selection for Performance Evaluation.....	81
6.3.2	The Result of Performance Evaluation on Sparsity Problem.....	84
6.3.3	The Result of Performance Evaluation on Scalability Problem.....	86
6.3.4	The Result of Performance Evaluation on Cold-start Problem.....	87
6.3.5	Overall Performance Evaluation.....	89

CHAPTER VII CONCLUSION AND FUTURE WORK.....	93
7.1 Conclusion.....	93
7.2 Future Works	94
REFERENCES.....	95
APPENDICES.....	97