

LIST OF CONTENTS

APPROVAL PAGE	i
STATEMENT	ii
LIST OF CONTENTS	iv
LIST OF FIGURES	viii
LIST OF TABLES	xii
LIST OF EQUATIONS	xiii
ABSTRACT	xiv
CHAPTER I: INTRODUCTION	1
1.1 Research Background	1
1.2 Research Problem	6
1.3 Research Scope	6
1.4 Research Objective	7
1.5 Research Advantage	7
CHAPTER II: LITERATURE REVIEW	9
CHAPTER III: THEORETICAL BASIS	16
3.1 Continuous Variable Prediction & Clustering	16
3.2 Analytical System	17
3.2.1 Data Transformation and Processing	17
3.2.2 Data Visualization	19
3.3 Pearson Correlation	19
3.4 Random Forest Regression	20
3.5 Hyperparameter Tuning	21

3.6 K-Nearest Neighbors Regression	22
3.7 K-Means Clustering	23
3.8 Elbow Method	25
3.10 Performance Metrics	26
3.10.1 Mean Squared Error	26
3.10.2 Mean Absolute Error	26
3.10.3 Root Mean Squared Error	27
3.10.4 R-Squared	27
CHAPTER IV: RESEARCH METHODOLOGY	29
4.1 Research Description	29
4.2 Tools and Materials	30
4.2.1 Tools	30
4.2.2 Materials	31
4.3 Literature Study	31
4.4 Data Acquisition	31
4.5 Preprocessing and Subsetting	33
4.6 Feature Analysis and Extraction	33
4.7 Data Normalization	35
4.8 K-Means Clustering Phase	36
4.9 Model Development for Rating Prediction	37
4.10 Output Representation and Evaluation	40
CHAPTER V: IMPLEMENTATION	42
5.1 Data Preprocessing and Encoding	42

5.2 Rating Prediction Model Building	51
5.3 Implementation and Evaluation of Rating Prediction Model (All Apps)	52
5.3.1 Training Phase - Lavanya Dataset	53
5.3.2 Testing and Evaluation - Lavanya Dataset	54
5.4 Implementation and Evaluation of Rating Prediction Model (Only Games)	55
5.3.1 Training Phase - Lavanya Dataset	58
5.3.2 Testing and Evaluation - Lavanya Dataset	59
5.5 Feature Analysis using K-Means Clustering	59
5.6 Exploratory Data Analysis	64
CHAPTER VI: RESULTS AND ANALYSIS	71
6.1 Performance and Experiment Summary of Rating Prediction Model	71
6.1.1 Random Forest Regressor Tuning	71
6.1.1.1 All apps model	71
6.1.1.2 Only games model	72
6.1.2 K-Nearest Neighbors (KNN) Regressor Tuning	74
6.1.2.1 All apps model	74
6.1.2.2 Only games model	75
6.1.3 Performance comparison	76
6.1.3.1 Improvement Experiment and Testing	79
6.2 Performance and Experiment Summary of K-Means Clustering	83
CHAPTER VII: CONCLUSION AND SUGGESTIONS	87
7.1 Conclusion	87

7.2 Suggestions	90
-----------------	----

REFERENCES	91
------------	----