UNIVERSITAS
GADJAH MADA
BOOK HYBRID RECOMMENDATION SYSTEM BASED ON MATRIX FACTORIZATION & METADATA EMBEDDING
TEGAR TAUFIK RAHMAN, Edi Winarko, Drs., M.Sc., Ph.D
Universitas Gadjah Mada, 2022 | Diunduh dari http://etd.repository.ugm.ac.id/

CONTENTS

COVER i
APPROVAL PAGE ii
STATEMENT PAGE iii
FOREWORD iv
CONTENTS v
LIST OF FIGURES viii
LIST OF TABLE x
INTISARI xi
ABSTRACT xii
CHAPTER I INTRODUCTION 1
1.1 Research Background 1
1.2 Research Problem 3
1.3 Research Objective 3
1.4 Research Scope 4
1.5 Research Advantage 4
CHAPTER II LITERATURE REVIEW 5
CHAPTER III THEORETICAL BASIS 16
3.1 Recommendation System 16
3.2 Content Based Filtering 16
3.3 Collaborative Filtering 18
3.4 Hybrid Recommendation System 21
3.5 Matrix Factorization 21
3.6 Hybrid Matrix Factorization 22
3.7 Feature Embedding 24
3.8 Term Frequency – Inverse Document Frequency 24
3.9 Loss Functions 25
3.9.1. Logistic Loss Functions 25
3.9.2. Bayesian Personalized Ranking Pairwise Loss 26
3.9.3. Weighted Approximate Rank Pairwise Loss 26
3.10 Evaluation Metrics 26
3.10.1 Receiver Operating Characteristics 26
3.10.2 Area Under Curve 27
3.10.3 Mean Reciprocal Rank 28

UNIVERSITAS
GADJAH MADA
BOOK HYBRID RECOMMENDATION SYSTEM BASED ON MATRIX FACTORIZATION & METADATA EMBEDDING
TEGAR TAUFIK RAHMAN, Edi Winarko, Drs., M.Sc., Ph.D
Universitas Gadjah Mada, 2022 | Diunduh dari http://etd.repository.ugm.ac.id/

CHAPTER IV RESEARCH METHODOLOGY 29
4.1 Research Description 29
4.2 Research Phase 29
4.3 Data Collection 29
4.4 Data Preprocessing 31
4.5 Feature Engineering 34
4.6 Modelling 35
4.7 Model Evaluation 39
4.8 Generate Recommendation 40
CHAPTER V IMPLEMENTATION 42
5.1 Tools and Materials 42
5.2 Data Collection 42
5.3 Data Pre-processing 42
5.3.1 Interaction datasets creation 43
5.3.2 Books Datasets Creation 44
5.3.3 User Datasets Creation 44
5.3.4 Selection of top 60 categories that appear in the book dataset 45
5.3.5 Selection of the top 60 categories that appear in the user dataset 46
5.3.6 Category column merge 47
5.3.7 Initializing the book’s category reference with 0 47
5.3.8 Creating user’s average and total rating columns 48
5.3.9 Creating book’s average and total rating columns 49
5.3.10 Converting rating into 1 and -1 49
5.3.11 Clean book and user data skewness 50
5.4 Feature Engineering 50
5.4.1 Creation of user’s category tf-idf dataframe 51
5.4.2 Creation of book’s category tf-idf dataframe 51
5.4.3 Initializing user’s category preference with 0 52
5.4.4 Creation of user book’s category preference 52
5.4.5 Creation of book’s category reference 53
5.5 Modelling 54
5.5.1 Collaborative Filtering Model 55
5.5.2 Hybrid Recommendation Model 56
5.6 Model Evaluation 61
5.6.1 AUC Score 61

UNIVERSITAS
GADJAH MADA
BOOK HYBRID RECOMMENDATION SYSTEM BASED ON MATRIX FACTORIZATION & METADATA EMBEDDING
TEGAR TAUFIK RAHMAN, Edi Winarko, Drs., M.Sc., Ph.D
Universitas Gadjah Mada, 2022 | Diunduh dari http://etd.repository.ugm.ac.id/

5.6.2 Reciprocal Rank 61
5.7 Generate Recommendations. 62
CHAPTER VI RESULTS AND DISCUSSIONS 68
6.1 Data Preprocessing Results 68
6.2 Feature Engineering Results 68
6.3 Experiment Results 70
6.3.1 Area Under Curve 70
6.3.2 Reciprocal Rank 73
6.3.3 Generate Recommendation Results 76
CHAPTER VII CONCLUSION AND SUGGESTION 79
7.1 Conclusion 79
7.2 Suggestions 79
REFERENCE 80