



**PROJECT REPORT
IMPLEMENTATION LEARNING VECTOR
QUANTIZATION (LVQ) ALGORITHM TO
CLASSIFY BOOK CATEGORIES BASED ON
BOOK'S DESCRIPTION**

**AMALIA ISMARDIANI
14.K1.0081**

**Faculty of Computer Science
Soegijapranata Catholic University
2020**

APPROVAL AND RATIFICATION PAGE



Final Project Title : Implementation Learning Vector
Quantization (LVQ) Algorithm to
Classify Book Categories Based on
Book's Description

Submitted by : Amalia Ismardiani

ID : 14.K1.0081

Date approved : July 13, 2020

Approved by

Supervisor : Hironimus Leong S.Kom., M.Kom.

Examiners 1 : R. Setiawan Aji Nugroho S.T.,
MCompIT., Ph.D

Examiners 2 : Hironimus Leong S.Kom., M.Kom.

Examiners 3 : Rosita Herawati S.T., M.I.T.

Examiners 4 : Y.b. Dwi Setianto

Head of Undergraduate Program : Rosita Herawati S.T., M.I.T.

Dean of Faculty of Computer Science : R. Setiawan Aji Nugroho S.T.,
MCompIT., Ph.D

This page is a legitimate page and can be verified via the address below.

sintak.unika.ac.id/skripsi/verifikasi/?id=14.K1.0081

STATEMENT OF ORIGINALITY

I, the undersigned:

Name : AMALIA ISMARDIANI

ID : 14.K1.0081

Certify that this project was made by myself and not copy or plagiarize from other people, except that in writing expressed to the other article. If it is proven that this project was plagiarizes or copy the other, I am ready to accept a sanction.



Semarang, July, 13, 2020



AMALIA ISMARDIANI
14.K1.0081

**APPROVAL PAGE FOR PUBLICATION OF
SCIENTIFIC PAPERS FOR ACADEMIC INTEREST**

The undersigned below :

Name : Amalia Ismardiani
Undergraduate Program : Technical Information
Faculty : Computer Science
Type of work : Skripsi

Approved to give Non-Exclusive Royalty Free Right to Soegijapranata Catholic University Semarang for scientific work entitled "**Implementation Learning Vector Quantization (LVQ) Algorithm to Classify Book Categories Based on Book's Description**" along with the existing tools (if needed). With this Non-Exclusive Royalty Free Right Soegijapranata Catholic University has the right to store, transfer media / format, manage in the form of a database, maintain, and publish this final project as long as I keep my name as a writer / creator and as a Copyright owner.

This statement I made in truth.

Semarang, July 13 2020

Sincerely

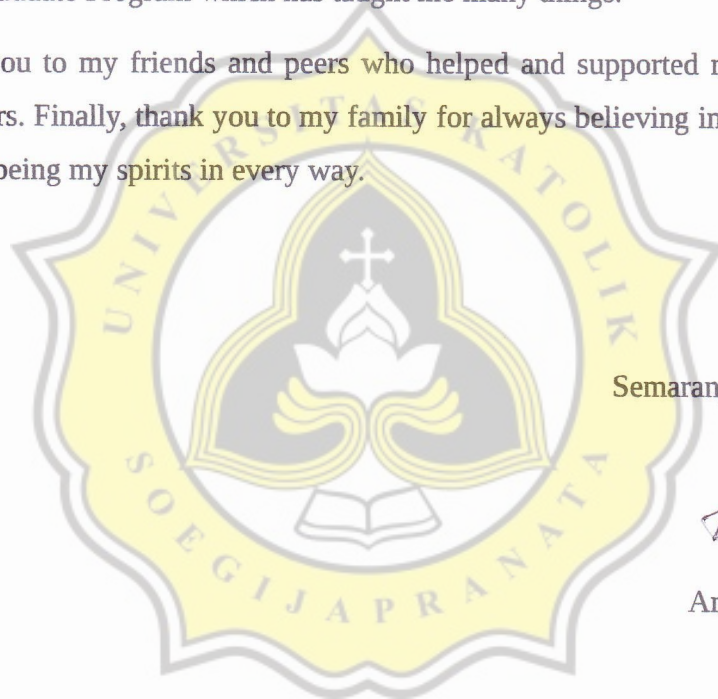


Amalia Ismardiani

ACKNOWLEDGEMENTS

Appreciation and thanks are given to Hironimus Leong S. Kom., M. Kom. as my Supervisor for helped me in completing this final project. I want to thank Rosita Herawati S.T., M.I.T. as my Academic Adviser for the help and support in completing my studies. I am also grateful to the Technical Information Undergraduate Program which has taught me many things.

Thank you to my friends and peers who helped and supported me during these past years. Finally, thank you to my family for always believing in me, supporting me and being my spirits in every way.



Semarang, July 13 2020

A handwritten signature in black ink, appearing to be 'Amalia Ismardiani'.

Amalia Ismardiani

ABSTRACT

In this era many bookstores use the internet to sell their books. But many bookstore websites show books that are not categorized hence the search is ineffective. To solve this problem by creating a program that can classify books based on book descriptions. Book categories can be determined based on book descriptions because books in a category can have the same description patterns

In this final project used implementation Learning Vector Quantization (LVQ) algorithm to the classification of book categories. The first step is to collect training and testing data using Web Scraper Google Chrome Extension. And then do text preprocessing methods to process the data and TF-IDF methods to extract data features. The accuracy of the LVQ algorithm is obtained by testing the learning rate and maxEpoch.

The testing shows that learning rate are important parameter to determine significant performance. The smaller the learning rate will resulting higher accuracy. The highest accuracy obtained at the end of the testing process ont the training data is 91.22%.

Keyword: Learning Vector Quantization, LVQ, Artificial Neural Network

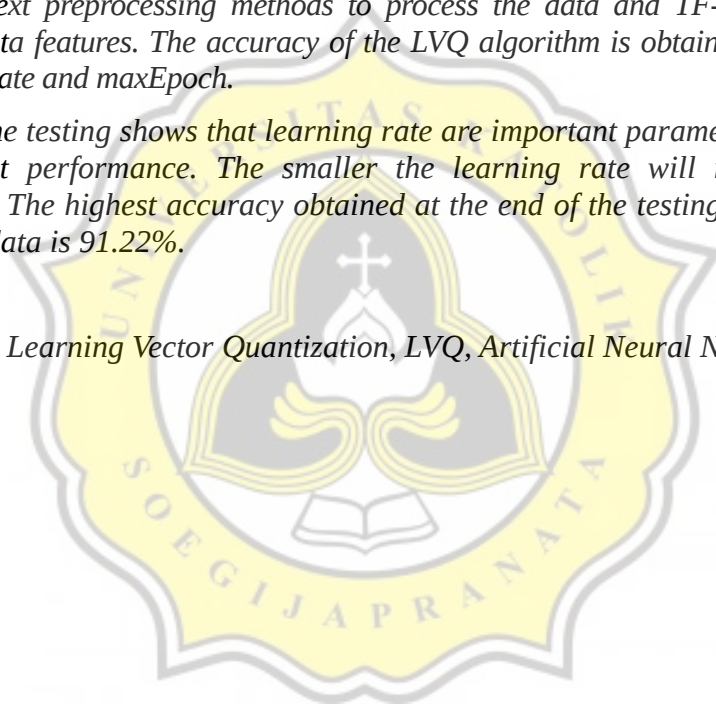


TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION.....	1
1.1 Background.....	1
1.2 Problem Formulation.....	2
1.3 Scope.....	2
1.4 Objective.....	2
CHAPTER 2 LITERATURE STUDY.....	3
CHAPTER 3 RESEARCH METHODOLOGY.....	8
3.1 Research Instrument.....	8
3.2 Data Collection.....	8
3.3 Research Procedure.....	8
CHAPTER 4 ANALYSIS AND DESIGN.....	11
4.1 Analysis.....	11
4.1.1 Collecting Data.....	11
4.1.2 Text Preprocessing.....	13
4.1.2.1 Case Folding.....	13
4.1.2.2 Tokenizing.....	14
4.1.2.3 Stop Word.....	14
4.1.2.4 Stemming.....	15
4.1.3 Term Frequency-Inverse Document Frequency (TF-IDF).....	15
4.1.4 Data Processing using LVQ Algorithm.....	17
4.1.4.1 LVQ ALGORITHM ANALYSIS ON TRAINING DATA.....	18
4.1.4.2 LVQ ALGORITHM ANALYSIS ON TESTING DATA.....	22
4.2 Design.....	24
4.2.1 Use Case Diagram.....	24
4.2.2 Flow Chart.....	25
4.2.3 UML Class Diagram.....	28
4.2.4 Database Schema.....	33
CHAPTER 5 IMPLEMENTATION AND TESTING.....	35
5.1 Implementation.....	35
5.1.1 Training Data.....	35
5.1.2 Testing Data.....	40
5.2 Testing.....	42
CHAPTER 6 CONCLUSION.....	44
6.1 Conclusion.....	44
6.2 Suggestions.....	44

ILLUSTRATION INDEX

Illustration 3.1: Program System Design.....	9
Illustration 4.1: Text Preprocessing Chart.....	13
Illustration 4.2: TF-IDF Result.....	17
Illustration 4.3: Use Case Diagram.....	24
Illustration 4.4: Training Data Flow Chart.....	25
Illustration 4.5: Testing Data Flow Chart.....	26
Illustration 4.6: LVQ Process Flow Chart.....	27
Illustration 4.7: Case Folding Training Class Diagram.....	28
Illustration 4.8: Case Folding Testing Class Diagram.....	28
Illustration 4.9: Tokenize Training Flow Class Diagram.....	29
Illustration 4.10: Tokenize Testing Flow Class Diagram.....	29
Illustration 4.11: Stop Word Training Class Diagram.....	30
Illustration 4.12: Stop Word Testing Class Diagram.....	30
Illustration 4.13: Stemming Training Class Diagram.....	31
Illustration 4.14: Stemming Testing Class Diagram.....	31
Illustration 4.15: TF-IDF Training Class Diagram.....	32
Illustration 4.16: Classification Class Diagram.....	32
Illustration 4.17: TfIdfTesting Class Diagram.....	33
Illustration 4.18: Training Data Tables in the Database.....	33
Illustration 4.19: Testing Data Tables in the Database.....	34
Illustration 5.1: Training Data Graphic Interface.....	35
Illustration 5.2: Testing Menu Graphic Interface.....	41

INDEX OF TABLES

Table 2.1: Accuracy Result.....	4
Table 4.1: Training Data Example.....	11
Table 4.2: Example of Case Folding Process.....	13
Table 4.3: Example of Tokenizing Process.....	14
Table 4.4: Example of Stop Word Process.....	15
Table 4.5: Example of Stemming Process.....	15
Table 4.6: Example of Training Data Representation.....	18
Table 5.1: LVQ Algorithm Results on Training Data.....	42
Table 5.2: Testing Results of LVQ Algorithm on Testing Data.....	43

