CLASSIFICATION TECHNIQUES FOR HANDWRITING DIFFICULTIES AMONG CHILDREN IN EARLY STAGE OF ACADEMIC LIFE

ANITH ADIBAH BINTI HASSEIM

A thesis submitted in fulfilment of the requirements for the award of the degree of Master of Engineering (Electrical)

Faculty of Electrical Engineering
Universiti Teknologi Malaysia

JUNE 2015
Specially dedicated to my beloved father and mother
Hasseim bin Shaaban and Robiah binti Romli
Also my beloved husband
Ismadi bin Ibrahim
My daughter Nur Auni Imthithal binti Ismadi
brothers, sisters and all my friends
for their inspiration, support and encouragement
throughout my adventure of educations
ACKNOWLEDGEMENT

First and foremost, I would like to express my grateful to Allah S.W.T because give me a good health and strength to complete my study. Without His permits, I would not be able to reach up until this level.

I wish to express my sincere appreciation to my project supervisor, Associate Professor Dr. Rubita binti Sudirman for her guidance and enthusiasm given throughout the progress of this study. Thanks for giving a worth knowledge that’s help to improve myself.

My appreciation also goes to my husband and family who have been so tolerant and supported me all these years. Thanks for their encouragement, love and emotional supports that they had given to me.

I would also like to thank Dr. Puspa Inayat binti Khalid for her co-operations, guidance and helps in this study.

Nevertheless, my great appreciation dedicated to all my friends and those whom involve directly or indirectly with this study. There is no such meaningful word than thank you so much.
ABSTRACT

In today's era, all aspects of complex occupational task, plus the importance of early identification of developmental disorders in children, demand the essential need for screening children’s handwriting at elementary schools. Many underlying competence structures may interfere with handwriting performance. Children starting their academic programme should be tested for their handwriting abilities and readiness through regular routine screening. Screening a vast majority of 4 to 7+ years old necessitate the use of automated systems to collect data, keep tracks, and increase the speed of analysis and accuracy. Based on Handwriting Proficiency Screening Questionnaire (HSPQ) evaluated by their teachers, 120 pupils were individually tested on their use of graphic production rules. Then, the samples were divided into two group of writers; below average writers (test group) and above average writers (control group) based on the score of HSPQ. Each participant was required to copy four basic lines in two opposite directions and trace a sequence of rotated semi circles. This research examines the dynamic features such as ratio of time taken and standard deviation of pen pressure. In this study, three classification methods: Artificial Neural Network, Logistic Regression and Support Vector Machine (SVM) were chosen to classify children with handwriting problem. 10-fold cross-validation method is used for testing and training. At the end of this study, the results among these classifiers and features were compared. Based on the results, it can be concluded that the performance of SVM with Radial Basis Function kernel is the best among classifiers as it gives 100% of screening accuracy.
ABSTRAK

TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECLARATION</td>
<td>i</td>
<td></td>
</tr>
<tr>
<td>DEDICATION</td>
<td>ii</td>
<td></td>
</tr>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td>iii</td>
<td></td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>iv</td>
<td></td>
</tr>
<tr>
<td>ABSTRAK</td>
<td>v</td>
<td></td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>vi</td>
<td></td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
<td></td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>xi</td>
<td></td>
</tr>
<tr>
<td>LIST OF APPENDICES</td>
<td>xii</td>
<td></td>
</tr>
</tbody>
</table>

1 INTRODUCTION
1.1 Overview 1
1.2 Problem Statement 3
1.3 Objectives of Study 4
1.4 Scopes of Study 4
1.5 Significant of Study 5
1.6 Thesis Organization 5
2 LITERATURE REVIEW

2.1 Introduction 7
2.2 Handwriting Development 7
2.3 Handwriting Difficulties 10
2.4 Dynamic Features 11
2.5 Classification Method 13
   2.5.1 Artificial Neural Network 14
   2.5.2 Logistic Regression 17
   2.5.3 Support Vector Machine 19
2.6 Cross Validation 22
2.7 Chapter Summary 23

3 METHODOLOGY

3.1 Introduction 25
3.2 Dataset 25
   3.2.1 Procedures 27
   3.2.2 Copying Task 28
   3.2.3 Tracing Task 29
3.3 Outcome Measure 30
   3.3.1 Copying Task 30
   3.3.2 Tracing Task 31
3.4 Features 32
3.5 Classifiers 33
   3.5.1 Architecture of Neural Network 34
   3.5.2 Architecture of Support Vector Machine 37
   3.5.3 Architecture of Logistic Regression 38
3.6 Chapter Summary 39
## 4 RESULTS AND DISCUSSIONS

4.1 Introduction 41
4.2 Classification Performance 41
  4.2.1 Artificial Neural Network Classification 42
  4.2.2 Support Vector Machine Classification 43
  4.2.3 Logistic Regression Classification 47
4.3 Combining Features 49
4.4 Chapter Summary 54

## 5 CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions 55
5.2 Recommendations for Future Works 57

REFERENCES 58

Appendices A-B 63
### LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Dynamic features of poor writers</td>
<td>12</td>
</tr>
<tr>
<td>3.1</td>
<td>Neural network training parameters</td>
<td>34</td>
</tr>
<tr>
<td>3.2</td>
<td>Samples distribution using 10-folds cross validation technique</td>
<td>35</td>
</tr>
<tr>
<td>4.1</td>
<td>Accuracy of prediction based on ANN</td>
<td>42</td>
</tr>
<tr>
<td>4.2</td>
<td>The classification performance of ANN</td>
<td>43</td>
</tr>
<tr>
<td>4.3</td>
<td>Accuracy of prediction based on linear SVM</td>
<td>44</td>
</tr>
<tr>
<td>4.4</td>
<td>Accuracy of prediction based on SVM with RBF kernel</td>
<td>45</td>
</tr>
<tr>
<td>4.5</td>
<td>Accuracy of prediction based on SVM with polynomial kernel</td>
<td>46</td>
</tr>
<tr>
<td>4.6</td>
<td>Accuracy of prediction based on LR</td>
<td>47</td>
</tr>
<tr>
<td>4.7</td>
<td>The classification results for ANN, LR and SVM based on combining features</td>
<td>50</td>
</tr>
<tr>
<td>4.8</td>
<td>Results of classification on children's handwriting</td>
<td>52</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>A simple neural network diagram</td>
<td>15</td>
</tr>
<tr>
<td>2.2</td>
<td>Sigmoid function</td>
<td>16</td>
</tr>
<tr>
<td>2.3</td>
<td>An ideal graph of sigmoid logistic function</td>
<td>18</td>
</tr>
<tr>
<td>2.4</td>
<td>Optimal separating hyperplane</td>
<td>20</td>
</tr>
<tr>
<td>2.5</td>
<td>Procedure of 10-fold cross validation</td>
<td>23</td>
</tr>
<tr>
<td>3.1</td>
<td>The digitizing graphic tablet</td>
<td>28</td>
</tr>
<tr>
<td>3.2</td>
<td>A notion of eight directions</td>
<td>29</td>
</tr>
<tr>
<td>3.3</td>
<td>Tracing task given to the participants</td>
<td>29</td>
</tr>
<tr>
<td>3.4</td>
<td>Percentage of participants constructing a sequence of semicircles in non-preferential direction</td>
<td>32</td>
</tr>
<tr>
<td>3.5</td>
<td>Flowchart illustrating neural network training process</td>
<td>36</td>
</tr>
<tr>
<td>3.6</td>
<td>SVM flowchart</td>
<td>38</td>
</tr>
<tr>
<td>4.1</td>
<td>The classification results for ANN, SVM and LR</td>
<td>48</td>
</tr>
<tr>
<td>4.2</td>
<td>Overall classification results</td>
<td>53</td>
</tr>
</tbody>
</table>
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>HL</td>
<td>Horizontal Leftward</td>
</tr>
<tr>
<td>HR</td>
<td>Horizontal Rightward</td>
</tr>
<tr>
<td>HSPQ</td>
<td>Handwriting Proficiency Screening Questionnaire</td>
</tr>
<tr>
<td>LD</td>
<td>Left Oblique Downward</td>
</tr>
<tr>
<td>LR</td>
<td>Logistic Regression</td>
</tr>
<tr>
<td>LU</td>
<td>Left Oblique Upward</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean Square Error</td>
</tr>
<tr>
<td>RBF</td>
<td>Radial Basis Function</td>
</tr>
<tr>
<td>RD</td>
<td>Right Oblique Downward</td>
</tr>
<tr>
<td>RU</td>
<td>Right Oblique Upward</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>VD</td>
<td>Vertical Downward</td>
</tr>
<tr>
<td>VU</td>
<td>Vertical Upward</td>
</tr>
</tbody>
</table>
### LIST OF APPENDICES

<table>
<thead>
<tr>
<th>APPENDIX</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>List of Publications</td>
<td>63</td>
</tr>
<tr>
<td>B</td>
<td>Handwriting Proficiency Screening Questionnaire (HSPQ)</td>
<td>64</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 Overview

All aspects of complex occupational task as well as the importance of early identification of developmental problems in children, arises the importance to screen children’s handwriting at elementary schools. On the development of digitalized modern world today such as keyboard with its fast typing task, can never be substitute handwriting. When writing or drawing, children perceptually extract relationships and perform according to their ability and skills. Many underlying competence structures may interfere with handwriting performance.

Penmanship must be a process associated with joy, esprit and self expression; instead of pressure, boring and complicated job for beginners. Academics argue handwriting is more than the transcription of ideas; it is directly related to how people generate and process those ideas. Handwriting is a language act, rather than just a motor act. It is about training the memory and hand to work together to correct and generate mental images and patterns of letters then interpret these into motor patterns of letters, automatically and without effort (Bloom, 2008). Reports emphasize that by ignoring poor handwriting may leads to fails to address a significant and continuing barrier to children's’ achievement. The average or poor handwriting children have only 40% chance of achieving level 4, in national tests (Bloom, 2008); which indicates that bad handwriting can lead to unsuccessful in the
examination by 40%. Therefore children need to develop the sub skills for handwriting at a young age, improving their basic hand eye and motor coordination, balance and visual perception.

Children starting their academic program should be tested for their handwriting abilities and readiness the same way as they are tested for optometry and dentistry through regular routine screening. Screening a vast majority of 4 to 7+ year old children not only to increase the speed of analyzing and processing data but also to collect data and keep track and increase the accuracy. Additionally, it significantly helps psychologists to monitor improvements and come up with modified new standard assessments based on the real large database. Moreover, this system will truly give a great benefit to the educational system and consequently the whole society’s psychological and physical health.

Various softwares have been presented for handwriting recognition and movement analysis but, softwares directly related to child handwriting analysis with the prospective of screening children in general, and addressing difficulties are rare and the research is in its early stage. This research will hypothesize and examine the dynamic features that used by Khalid (Khalid, et al., 2010(a, b)). In contrast to Khalid approach (Khalid, 2012), different technique will be used to measure the influence of each feature in classifying pupil’s handwriting performance. In this project, we propose using three classification methods which are Artificial Neural Network (ANN), Logistic Regression (LR) and Support Vector Machine (SVM).
1.2 Problem Statement

Writing is an important development skill for a child to master even though computers are widely used nowadays. Several study have been done to analyze pupils with handwriting difficulties, however most of the studies that involve in handwriting movement only give an attention to children with known physical or psychological problem. Nevertheless, not all these problems can be categorized as clear cut disease and condition. Hence, an effective solution should be identifying to indicate pupils who have difficulties in writing.

Several classifiers had been used in the literature reviews (Khalid, 2012; Guest et. al., 2003; Chindaro et. al., 2004) such as Hidden Marcov Model, Artificial Neural Network and Logistic Regression to select those who have handwriting problem. However, the maximum classification accuracy of these classifiers is just around 83% which is not highly enough to highlight the behavior differences between average and below average writers. Thus, it is the attention of this study to pilot the objective and select the best classification technique to increase the accuracy of prediction.

In contrast with similar method known by Khalid (2012), each different feature was tested individually and this study describe experiments carried out using Support Vector Machine (SVM) in addition to those classification methods used in previous researches. SVM is a supervised learning method that has proven it’s efficiently over classic Neural Networks and its subset (Burges, 1998). The advantages of SVM are good generalization performance, able to handle high dimensional data and able to map the data into new high dimensional feature space for better classification using kernel functions.
1.3 **Objectives of Study**

The objectives of this study are:

i. To compare three classification methods to determine pupils who have difficulties in writing.

ii. To classify children either at risk of handwriting difficulties or not based on drawing task.

1.4 **Scope of Study**

The scope of the study is used as the guideline of the study. In order to achieve the objectives, the scope of the project has been confined as follows:

i. Data collection: normal healthy children between 7 to 12 years old in Skudai district performing copying and tracing tasks.

ii. Apply the different types of classifiers which are Artificial Neural Network, Logistic Regression and Support Vector Machine.

iii. Analyse the parameters based on the collected data.

iv. The standard deviation of pen pressure, ratio of time taken and the used of progression rules are the significant features that were used to identify children who are at risk of handwriting difficulties.
1.5 Significant of Study

This study investigates handwriting performance of a normal children aged between 7 and 12 years old. From the results, it may be helpful to the teachers because it can serve as a guide to deal with the problems and topic related to handwriting difficulties. The teachers also can monitor the students who have been identified to have difficulties in writing and enable them to plan an action based on instructional programme that suited to the students’ strengths and weaknesses.

Moreover, this study significantly helps psychologists to monitor improvements and come up with modified new standard assessments based on the real large database. On the other hand, this system will truly give a great benefit to the educational system and consequently the whole society’s psychological and physical health.

The student will be the most benefited by the results of this study. The finding will guide them in terms of their limitation and weaknesses. In addition, the students who have difficulties in handwriting can have more attention on the writing’s physical process, thus limiting use of higher order cognitive planning, skills and generation of content. Therefore, the students will manage to write and complete their writing task efficiently.

1.6 Thesis Organization

This thesis consists of five chapters. Chapter 1 introduces the background of the research, problem statement, objective of the study, scope of the study and the overall thesis outline.

Chapter 2 focuses on the handwriting problems as well as the development of handwriting project. This chapter also give the explanation about dynamic features
that used in this study. Furthermore, the classification methods which are artificial neural network, support vector machine and logistic regression have been discussed at the end of this chapter.

Chapter 3 reports the details of the dataset used in this study. It includes the process of developing the supplementary screening as well as extraction of dynamic features. Besides, this chapter also discusses the implementation of three classification methods used in this study.

Chapter 4 explains and discusses the results obtained and the analysis made. Comparisons between the classifiers’ results are made in order to achieve the final objective of this study. Among these classifiers, further analysis has done with increasing the number of features for the best classifier only.

Finally, Chapter 5 gives the conclusion of this study. Last but not least, this chapter also gives some recommendation on future development of this study.
REFERENCES


models in predicting violent re-offending". Journal of Qualitative Criminology, 27(4): 547-573.


