

Application Of Learning Vector Quantization Algorithm In Educational Games Learning To Write Aksara Bali

Agus Wisnawa^{a1}, I Gede Arta Wibawa^{a2}, Ida Bagus Gede Dwidasmara^{a3}, Made Agung Raharja^{a4}, I Komang Ari Mogi^{a5}, Luh Arida Ayu Rahning Putri^{a6}

^aInformatics Department, Faculty of Mathematics and Natural Sciences, Udayana University
South Kuta, Badung, Bali, Indonesia

¹haguswisnawa03@gmail.com

²gede.arta@unud.ac.id

³dwidasmara@unud.ac.id

⁴made.agung@unud.ac.id

⁵arimogi@unud.ac.id

⁶rahningputri@unud.ac.id

Abstract

Almost every region in the nusantara has its own regional script, one of which is in Bali the people know the Balinese script as a script commonly used by their ancestors in writing. In order to keep this Balinese script sustainable and not extinct, learning about this script is applied in schools in Balinese language subjects. In addition to applying learning about script in schools, it would be better if learning about Balinese script could be done in more interesting ways and concepts. One of them is by making an educational game for the introduction of Balinese characters. Which in this application the user will provide input in the form of Balinese script characters written on the application, which then the input will go through a preprocessing process then proceed with diagonal feature extraction, then the results of the feature extraction will go through a classification process using the Learning Vector Quantization method. . The result is a web-based application that can recognize Balinese script writing using the Learning Vector Quantization classification method with an accuracy rate of 58.6% and get a good response from each respondent who has tested the application.

Keywords: Balinese Script, Game Application, Diagonal Feature Extraction, LVQ, Classification.

1. Introduction

Balinese script is a regional writing originating from Bali. Balinese script has been used since ancient times until now. Balinese script can be used in writing Balinese and Sanskrit in various fields such as in religious ceremonies, arts, writing Balinese literature and so on. Therefore, the Balinese script must be preserved so that it is not lost and forgotten by the younger generation. One way to preserve Balinese culture, especially in learning Balinese script, is to teach the younger generation from an early age to learn Balinese script [1].

Based on a digital report from Hootsuite We are Social Indonesian in January 2021, the percentage of Indonesian internet users aged 16 to 64 who play games on a desktop / computer reaches 53% [2]. This indicates that the game on the computer is quite popular with the people of Indonesia. Games can also be used as learning media. Based on this, the researcher aims to create an educational game with the topic of Balinese script.

In the manufacturing process, a classification stage is needed to recognize the written characters. The Learning Vector Quantization (LVQ) method was chosen as the classification method because it has high accuracy, it is based on research from Solichin and Rahman who used the LVQ method as vehicle number plate identification with an accuracy of 95.32% [3].

2. Research Method

2.1 System Analysis

Classification is a process that aims to determine an object into a pre-determined class or category [4]. Before doing the classification there is a process that must be done, the process consists of preprocessing, feature extraction, and classification. The feature extraction method used in this study is diagonal feature extraction. Classification of data in this study using the method of Learning Vector Quantization (LVQ). In this study, the process of introducing Balinese script starts from the input image obtained from application users. This input image will then undergo preprocessing. After getting the preprocessing results, then proceed with the feature extraction process to get the class weight of the image. After getting the weights for each class, the classification process is carried out using the LVQ method. This classification process is divided into two, namely training or also called training and testing or can be called testing. The classification process aims to determine whether the input from the user is in accordance with the given command. After the classification process is complete, this process will end. The system analysis flowchart can be seen in Figure 1.

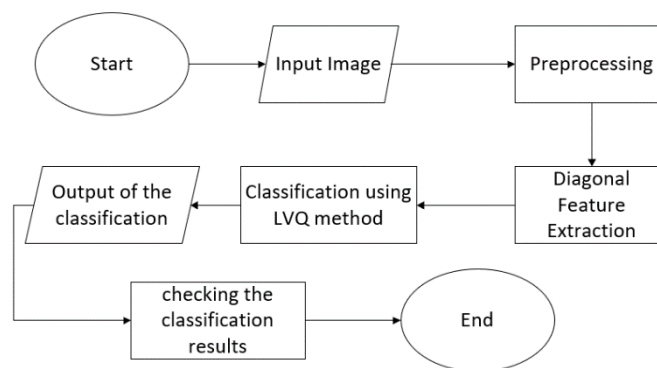


Figure 1 System Analysis Flowchart

2.2 Data

The wreastra script consists of 18 characters, namely ha, na, ca, ra, ka, da, ta, sa, wa, la, ma, ga, ba, nga, pa, ja, ya, nya [5]. The 18 characters are consonants, while the vowels are taken from the wisarga script, which is added by a pengangge aksara suara. The 6 pengangge aksara suara are ulu, suku, taling, tedong, taling tedong, and pepet [6]. The data used in this study is divided into two types, namely training data and test data. Training data is the data used to train the architecture of the LVQ method. The training data used in this study were primary. Students are asked to write each character on the provided paper 9 times, so that the total training data used is 90 pieces of writing on each Balinese script character. The details of the characters used are the characters in the Wreastra script, totaling 18 characters and the voice character cast, totaling 5 characters so that the total characters are 23 characters so that the total of all training data obtained is 2070 training data.

2.3 Preprocessing

In the preprocessing the initial data processing is divided into 2 processes, namely the process of processing test data and training data where both these processes have the same stages but use different data. This process consists of binaryzation, normalization, and thinning.

2.4 Feature Extraction

Diagonal feature extraction is one method that can be used to find the characteristics of handwritten characters. The first thing to do is to divide the input image into several zones that have the same size. Diagonal feature extraction is one method that can be used to find the characteristics of handwritten characters. The first thing to do is to divide the input image into several zones that have the same size. The results of the preprocessing are broken down into a size of 10x10 pixels so that

the image has 54 areas or features. In each of these features a diagonal line will be drawn according to the size of the feature, where with the size of each feature of 10x10 pixels it will produce 19 diagonal lines. the value of all foreground pixels in the feature. These values will be averaged to form a single value for the features they occupy. This process will be repeated until all features are scored.

2.5 Learning Vector Quantization (LVQ)

The Learning Vector Quantization (LVQ) algorithm has the result in the form of an appropriate weight value for grouping vectors into destination classes that have been initialized during the formation of the LVQ network. The testing algorithm of LVQ aims to calculate the output value (vector class) closest to the input vector, or can be equated with the classification process [7]. The LVQ training algorithm is as follows:

1. Initialization of the value of learning rate (α), learning rate deduction ($dec \alpha$), training data, maximum iteration and minimum α used.
2. For each input do steps a to d:
 - a. Calculate the euclidean distance of the input data with the weight vector for each class with Equation :

$$D(x, w) = \|X_i - W_j\| \dots \dots \dots (1)$$
 - b. Determine the minimum distance from the distance calculation results so that the output C_j is obtained.

$$C_j = \text{Min } D(x, w) \dots \dots \dots (2)$$
 - c. Fix the weights with the condition:

If $T = C$

$$W_j(\text{new}) = W_j(\text{old}) + \alpha[X_i + W_j(\text{old})] \dots \dots \dots (3)$$

If $T \neq C$

$$W_j(\text{new}) = W_j(\text{old}) - \alpha[X_i + W_j(\text{old})] \dots \dots \dots (4)$$
 - d. Make improvements to the value of learning rate (α)

$$\alpha(\text{new}) = \alpha(\text{old}) \times dec \alpha \dots \dots \dots (5)$$
3. Stops when the iteration is maximum or the learning rate is minimum.

Where:

- W : Vector Weight
- D : Distance data to i
- X : Data
- T : Target
- (α) : Learning rate

There are 3 layers of the LVQ architecture created, namely the input layer, the competitive layer, and the output layer. The input layer in this study is an input image consisting of 69 features. In the competitive and output layers, the architecture of this research consists of 23 classes used. The Learning Vector Quantization architecture in this study can be seen in Figure 2.

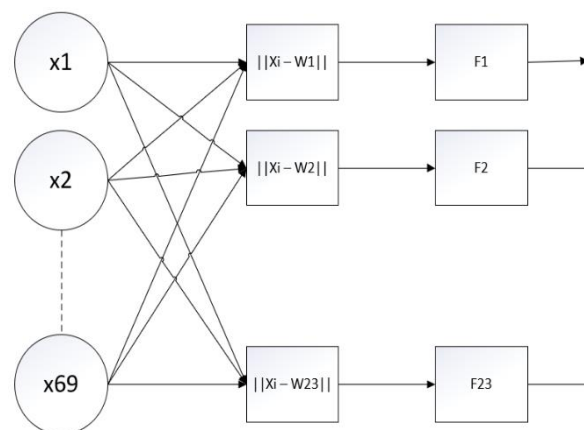


Figure 2. LVQ architecture

2.6 Testing

Testing process will be divided into 3 types, black box testing, accuracy test, and UAT (User Acceptance Test).

- a. Black Box Testing
Black box testing is a test that aims to show the function of the software about how it operates, whether the input and output have run as expected [8].
- b. Accuracy Testing
The level of character recognition accuracy is obtained by dividing the number of correct recognitions by all test results, the final result obtained will be in the form of a percentage [9]. As for this study, accuracy testing is carried out per character that appears and not per sentence. The results of this test are to determine the ability of the system in reading Balinese script for the Balinese script introduction game.
- c. User Acceptance Test (UAT)
User Acceptance Test or UAT is a test intended to determine user ratings of the applications made. The process of UAT testing is that prospective users are asked to use the application created and then the user is asked to fill out a questionnaire containing a perspective assessment of the application made [10].

3. Result and Discussion

3.1. Software Implementation

Interface Implementation Implementation of the interface on this system will be divided into 10 pages, namely the main menu page which can be seen in Figure 3, this page is the main page of the application. Next is the how to play page which can be seen in Figure 4, this page describes how to play. The next page about the application can be seen in Figure 5, this page contains information about the game being built such as the algorithm used, the technology to build the application, and also the creator of the application. The start page of the game which can be seen in Figure 6, this page becomes the second menu as an introduction for the user to the learning page or the practice page. On the learning page which can be seen in Figure 7, in this page users can see Balinese script and the latin writing. On the exercise page which can be seen in Figure 8, the user can practice the user's ability to write Balinese script which will be checked whether it is right or wrong. Furthermore, the correct answer page can be seen in Figure 9 or the wrong answer page which can be seen in Figure 10 is the page that is displayed after the user submits an answer on the exercise page.



Figure 3. Main Menu



Figure 4. How to Play



Figure 5. About Application



Figure 6. Game Start

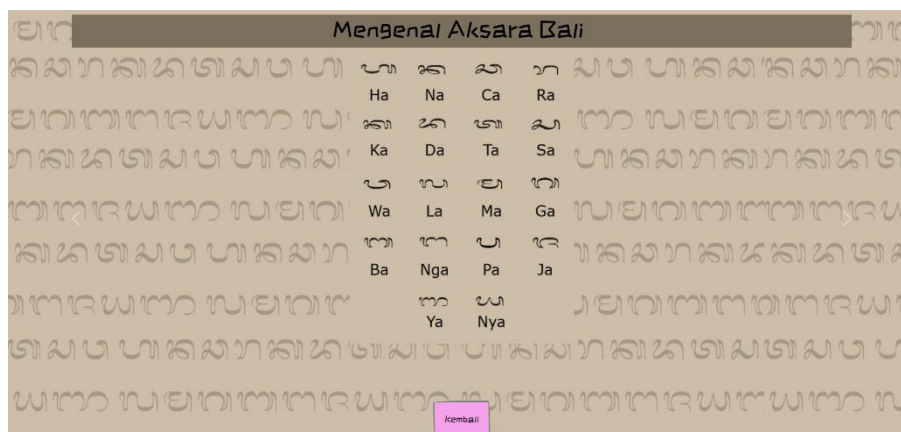


Figure 7. Study

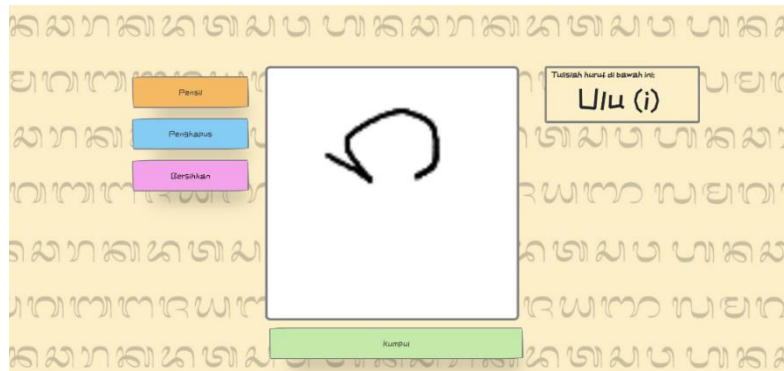


Figure 8. Practice



Figure 9. Correct Answer



Figure 10. Wrong Answer

3.2. Black Box Testing

This test will test the functionality of the application whether it is in accordance with the expected expectations. The results of the black box testing are in the table 1.

Table 1 Black Box Testing

No	Testing	Result	Description
1.	Start the app	Success	The application can be opened and display the main page.
2.	Press the "Start Game" button	Success	The button can direct the application to the game mode select page.
3.	Pressing the "How to Play" button	Success	The button can direct the application to the how-to page.
4.	Pressing the "About" button	Success	The button can direct the app to the about page.
5.	Pressing the "Learn" button	Success	The button can direct the application to the study page.
6.	Pressing the "Practice" button	Success	The button can direct the app to the workout page.
7.	Writing on canvas	Success	Users can write on the provided canvas.
8.	Pressing the "Eraser" button	Success	The user enters erase mode and can delete the existing doodles in the canvas manually.
9.	Pressing the "Pencil" button	Success	Users can return to writing mode as usual after previously using erase mode.
10.	Pressing the "Clean" button	Success	Users can clean the entire canvas content automatically.
11.	Press the "Enter" button	Success	This button will stop writing activities and then send data to enter the classification process.
12.	Pressing the "Done" button	Success	This button will stop the learning process on the study menu and return to the previous page.
13.	Question	Success	Questions can appear in the practice menu.
14.	Press the "Back to Start Page" button	Success	This button will direct the application to the start page.

3.3. Accuracy Testing

From the tests carried out, where each character was tested 10 times so that the total tests carried out were 230 tests with a total of 135 trials which were true and 95 were false. It was found that the accuracy of the classification of Balinese characters using the Learning Vector Quantization method with diagonal feature extraction is 58.6%.

3.4. UAT Testing

accuracy, next requires the calculation of metrics which is used to better present the performance of the model that has been created. Among them is Precision, which calculates how much data is really positive from all data that is predicted as positive, or in other words true positive, Recall calculates how much positive data is predicted as positive, then the last is the F-1 Score, which is a combination of Precision and Recall values to compare different models or commonly referred to as harmonics means of the two. The UAT test result can be seen in table 2.

Table 2 UAT Test Result

No.	Question	Answer					Score	Percentage
		A	B	C	D	E		
1.	Does the application look attractive?	0	7	3	0	0	37	74%
2.	Is the menu of the application easy to understand?	8	1	1	0	0	47	94%
3.	Is this app easy to use?	0	5	3	0	2	31	62%
4.	Can an example of Balinese script help you to know the shape of the Balinese script?	8	1	1	0	0	47	94%
5.	Can practice help you to know your ability to write Balinese script?	0	3	7	0	0	33	66
6.	Can this application help you to improve your writing skills in Balinese script?	0	1	8	1	0	30	60%
7.	Can this application make learning to write Balinese script fun?	7	2	1	0	0	46	92%
8.	Can this application be a good medium for learning Balinese script?	0	0	8	1	1	27	54%

Based on the UAT testing conducted, it can be concluded that the respondents agree that the application display is attractive. Furthermore, respondents strongly agree that the application menu is easy to understand. Respondents agree that the application is easy to use. Furthermore, respondents strongly agree that the existence of script samples can help to determine the shape of the Balinese script. Respondents agreed that the exercise can help respondents know the respondent's ability to write Balinese script. Furthermore, respondents agreed that this application can help respondents in improving their writing skills in Balinese script. Respondents strongly agree that this application can make the experience of learning to write Balinese script enjoyable. Finally, the respondents considered that this application could be a good medium for learning Balinese script. The total points

from all the questions obtained were 298 points or 74.5%. So it can be said that the application made has received a positive response from application users.

4. Conclusion

From the research that has been done, the results obtained are in the form of an educational game application that introduces Balinese script which is aimed at people, both students and the general public. The results of black box testing on this application show that all features and functions can run according to the desired expectations. Based on UAT testing, the application can be said to have received a positive response from the user.

This study also conducted an accuracy test. The purpose of this accuracy test is to measure the percentage of the ability of the LVQ (Learning Vector Quantization) algorithm with diagonal feature extraction in classifying Balinese script characters written by users. The results of this accuracy test are the percentage of success of the LVQ algorithm with diagonal feature extraction in classifying Balinese script writing, which is 58.6% in a total of 230 test trials.

References

- [1] P. P. G. Pertama, Suyoto And T. Suselo, "Pengembangan Aplikasi Mobile Pengenalan Aksara Bali Kendala Huruf Latin Dengan Augmented Reality," Seminar Nasional Teknologi Informasi Dan Komunikasi 2015(Sentika 2015), P. 1, 2015.
- [2] S. Kemp, "We Are Social Indonesian Digital Report," Hotsuite, 2021.
- [3] A. Solichin And Z. Rahman, "Identifikasi Plat Nomor Kendaraan Berbasis Mobile Dengan Metode Learning Vector Quantization," Jurnal Ticom Vol.3, 2015.
- [4] A. M. Puspitasari, D. E. Ratnawati And A. W. Widodo, "Klasifikasi Penyakit Gigi Dan Mulut Menggunakan Metode Support Vector Machine," Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer, Pp. 802-810, 2018.
- [5] I. Suasta, I. Mayun And W. Rupa, Modernisasi Dan Pelestarian Perkembangan Metode Dan Teknik Penulisan Aksara Bali, Jakarta: Proyek Pengkajian Dan Pembinaan Nilai-Nilai Budaya Direktorat Sejarah Dan Nilai Tradisional Direktorat Jenderal Kebudayaan Departemen Pendidikan Dan Kebudayaan , 1996.
- [6] M. I. K. Narwadha, Sareng Mapuruk Aksara Bali, Surabaya: Parāmita, 2006.
- [7] R. Meliawati, O. Soesanto And D. Kartini, "Penerapan Metode Learning Vector Qiantization (Lvq) Pada Prediksi Jurusan Di Sma Pgri 1 Banjarbaru," Kumpulan Jurnal Ilmu Komputer (Klik), 2016.
- [8] M. Adiwijaya, I. S And Y. Christyono, "Perancangan Game Edukasi Platform Belajar Matematika Berbasis Android Menggunakan Construct 2," Jurnal Teknik Elektro, Universitas Diponegoro Semarang, 2015.
- [9] R. Kohavi And F. Provost, Special Issue On Applications Of Machine Learning And The Knowledge Discovery Process, Kluwer Academic Publishers, 1998.
- [10] H. Supriyono, R. Rahmadzani, M. Adhantoro And A. Susilo, "Rancang Bangun Media Pembelajaran Dan Game Edukatif Pengenalan Aksara Jawa "Pandawa"," The 4th University Research Colloquium 2016, 2016.

This page is intentionally left blank