K-MEANS ALGORITHM TO DETERMINE THE EFFECTIVENESS OF USING QUIZIZZ IN ENGLISH LANGUAGE LEARNING FOR STUDENTS

MOEH ZAINAL KHAIRUL
Universitas Negeri Makassar
moehzainal@gmail.com

HARDIANTI HAFID
hardiantihf@unm.ac.id

Abstract

Mastery of the English language in the era of globalization has become an essential skill required in various aspects of life. English language proficiency is highly regarded, leading many individuals across the globe to opt for learning it as an additional language. Several countries have incorporated English into their school curriculum. Challenges in the teaching and learning of English, such as limited learning resources and unengaging teaching methods, call for innovative solutions to enhance learning effectiveness. This research explores the use of Quizizz, a game-based learning platform, as an interactive learning medium in teaching English to students at SMP IT Al-Akhyar Makassar. This study aims to assess the effectiveness of implementing Quizizz in improving students' understanding of English by using the K-Means algorithm for analyzing students' learning performance data. This study employed a quantitative research design and pre-experimental approach with pretest and posttest instruments for data collection. The research method involved collecting student scores before (pretest) and after (posttest) the implementation of Quizizz in English language learning. The data were analyzed using the K-Means algorithm to classify students into clusters based on their improvement in understanding. The evaluation of the clustering model was conducted by comparing the average within-cluster distance and the average between-cluster distance to assess the cohesiveness and differentiation among clusters. The results show a significant improvement in students' understanding of English after the implementation of Quizizz, indicated by the shift of student allocation into clusters with higher criteria, namely excellent. The evaluation of the clustering model shows that the clusters were formed effectively, with strong internal cohesion and clear differentiation between clusters. This supports the claim that Quizizz helps people learn English better.
Keywords: Quizizz, K-Means, Clustering, English Language

INTRODUCTION

Mastery of the English language has become an essential skill in today's era of globalization and information, required in various aspects of life, from education and business to international communication. Knowledge of English is considered very important, leading many individuals around the world to choose to learn this language as an additional language. Several countries have integrated English into their school curricula, allowing children to start learning English at an early age (Ilyosovna, 2024). English, as an international language, plays a crucial role in facilitating the exchange of knowledge and technology between countries and in accessing various global information sources. Indonesia, with its large population in a global environment that does not use English as the primary

language, increasingly shows its importance as a target market for education in English (Zein et al., 2020).

In Indonesia, the proportion of individuals who use English correctly and appropriately is still very minimal. Nevertheless, mastery of the English language can enhance Indonesia's reputation in the international community by efficiently facilitating the communication of this country's potential (Darma & Widiastuty, 2023). Indonesia needs to align itself with global developments as a country that is in the process of development, making the mastery of the English language an essential requirement.

Indonesia still faces a wide range of challenges in teaching and learning English, from limited learning resources and less engaging teaching methods to low motivation and interest among students. In Indonesia, learning English is one of the important aspects of the education system that requires an effective and innovative approach to enhance student understanding. The use of technology in the learning process has become a rapidly growing trend, with online educational applications becoming an increasingly popular solution. Classroom learning activities should incorporate the use of technology platforms. Various methods can be used to motivate students and teachers to seek new inspiration in the teaching and learning process through technology (Amalia, 2020). One popular platform is Quizizz. Quizizz is an enjoyable and interactive media in the educational field that allows students to engage in class activities with many participants (Zhao, 2019). Quizizz offers various features to enhance learning. It is one of the interactive tools that teachers can use in the classroom setting. This enjoyable learning media features colorful designs, and graphics made in the form of games, and users can choose avatar images (Khasanah & Lestari, 2021).

According to Adianto (2020) the collaboration of teaching methods in the classroom with Cooperative Learning, Scientific Approach, and Quizizz produces good results. The teaching steps carried out by teachers include: a) starting with greetings and singing the national anthem of Indonesia and checking student attendance, b) giving tasks that promote group-based solutions, c) group learning with the application of Quizizz in giving quizzes, making learning more interesting. Although Quizizz is promising as an effective learning tool, further research is needed to evaluate the extent of its effectiveness in enhancing students understanding of learning. The importance of evaluating the effectiveness of technology implementations like Quizizz in English language learning becomes a crucial aspect that cannot be overlooked.

In this research, the K-Means algorithm is one of the popular clustering methods in data mining and machine learning, offering analytical capabilities to group data based on certain characteristics, thus allowing teachers and researchers to assess and analyze the effectiveness of using Quizizz. Through the application of the K-Means algorithm, students can be grouped based on the scores they achieve by observing their improvement in understanding English, which can then be used to identify the effectiveness of the learning methods used. The K-Means algorithm is a suitable method for cluster analysis in data. (Syamfithriani et al., 2023). This study will use the K-Means algorithm to classify students based on their learning outcomes using Quizizz, allowing for a more in-depth evaluation of how much Quizizz has contributed to enhancing their understanding of English.

Therefore, this study aims to explore the application of the K-Means algorithm in determining the effectiveness of using Quizizz in English language learning among students at SMP IT Al-Akhyar Makassar, by analyzing student learning performance data collected through the Quizizz platform. This research strives to understand how technology can be effectively integrated into English language learning and how data analysis can assist in enhancing teaching strategies going forward. Through this approach, it is hoped that valuable insights can be gained into how technological innovations can be leveraged to improve the quality of English language learning, while also contributing to the development of more adaptive and responsive teaching methods to meet student needs.

METHODS

Research Design

The research design used in this study is a quantitative and pre-experimental method. Quantitative method is a systematic investigation into a relationship between measurable variables, typically on instruments, so that numerical data can be analyzed using statistical procedures (Creswell, 2014). In this study, the method employed utilizes K-Means algorithm statistical technique. Pre-experimental design applied is the one-group pretest-posttest design. In the one-group pretest-posttest design, there is a pretest stage before the treatment is administered. This is crucial as it enables comparison of conditions before and after the treatment, allowing for a more accurate evaluation of treatment outcomes (Sugiyono, 2013). Below is the One-Group Pretest-posttest design used in this research:

Table 1. Research Design of One-Group Pretest-Posttest

Pretest	Treatment	Posttest
T1	X	T2

T1 : Pretest to assess initial abilities

X : The utilization of Quizizz in English language learning

T2 : Posttest to assess final abilities

In learning evaluation, all types of tests and non-test instruments have their respective functions. Tests are classified based on their function as tools to measure students' learning progress (Matondang, 2009). Based on their form and type, instruments are divided into 2 categories, namely test instruments and non-test instruments (Magdalena et al., 2021). The tests used in this study consist of:

a. Pretest: An initial test is conducted using blocks as a medium to assess each child's level of creativity through four aspects of creativity (fluency, flexibility, originality, and elaboration). The pretest is conducted alternately in the teacher's room so that each child can comfortably take the test.
b. Treatment: The children are given treatment consisting of four Brain Gym movements that enhance creative thinking skills, namely the cross crawl, gravity glide, energy yawn, and hip circles. The treatment is carried out collectively using Quizizz-based learning.
c. Posttest: A final test similar to the pretest is conducted. This test is useful for observing differences before and after the treatment. The posttest is conducted alternately in the teacher's room.

Test instruments are used to measure the level of mastery of concepts so that students can enter the stage of creative thinking. To measure this mastery test, multiple-choice questions are used, created according to the indicators found in the learning materials, consisting of 25 questions with cognitive knowledge (P) and Skill (K) aspects as detailed in Tables 2, with the variable details as follows.

Table 2. Indicator of Pretest and Posttest Test Instrument

Indicator	The number of questions
P1: Identifying the expressions of asking likes and dislikes	5
P2: Collecting information about comparison	5
P3: Interpreting data in the form of comparison	5
K1: Creating the video of the conversation	5
K2: Creating conversation asking and giving direction	5

Data

Data collection was conducted using posttest and pretest instruments. The research population consists of a specific group to be studied with clear boundaries. The population for this study is the 7th-grade class at SMP IT Al-ahkyar in Makassar City, and the sample comprises 24 students. Sampling is the material that will be studied to ensure focused research. The sample selected for this study will be used as the experimental group using purposive sampling technique, which is a sampling technique based on specific considerations. Thus, primary data on student scores before and after the implementation of the Quizizz method in learning for 7th-grade students at SMP IT Al-ahkyar in Makassar City, South Sulawesi Province, were obtained. The results of the pretest and posttest conducted can be seen in Tables 3 and 4.

Table 3. Knowledge and Skills Aspect Scores Before the Application of Quizizz (Pretest)

Student Attendance Number Order	P1	P2	P3	K1	K2
1	96	95	90	96	95
2	90	88	87	90	88
3	88	89	90	88	89
4	90	87	88	90	87
5	89	87	88	89	87
6	95	96	89	95	96
7	95	94	92	95	94
8	90	93	96	90	93
9	87	90	92	89	90
10	89	90	95	89	90
11	89	87	90	89	87
12	87	89	90	87	89
13	90	93	95	90	93
14	95	89	80	95	89
15	87	86	89	87	86
16	89	90	87	89	90
17	93	90	87	93	90
18	89	87	86	89	87
19	98	94	92	98	94
20	98	87	93	98	87
21	93	94	95	93	94
22	87	89	84	87	89
23	89	90	92	86	90
24	80	94	85	80	94

Table 4. Knowledge and Skills Aspect Scores After the Application of Quizizz (Posttest)

Student Attendance Number Order	P1	P2	P3	K1	K2
1	98	100	95	97	89
2	90	95	100	90	88
3	88	100	90	88	89
4	90	87	88	90	87
5	100	90	99	90	87
6	100	100	94	95	90
7	95	94	92	95	94
8	100	95	94	97	94
9	87	90	92	89	90
10	95	97	89	100	89
11	89	87	90	89	87
12	100	95	100	95	89
13	100	98	99	95	90
14	95	90	100	100	90
15	87	86	90	87	86
16	89	91	90	100	90
17	93	100	87	93	90
18	89	87	86	89	87
19	98	94	100	98	94
20	100	100	95	98	90
21	100	100	90	97	94
22	87	89	84	87	89
23	89	90	92	86	90
24	80	94	85	80	94

Data analysis procedure

a. Collecting data using pretest and posttest instruments

b. Data Preprocessing

The data preprocessing technique is an important stage in data processing, carried out to avoid problems that can occur due to poor data quality. The data preprocessing stage used in this study is data integration, which involves combining multiple datasets into one, resulting in data as shown in Tables 1 and 2.

c. Implementation of K-Means

At this stage, K-Means is implemented to see the effectiveness of student scores before and after the application of Quizizz, with the following steps:

1) import data into the R application
2) determining the number of groups/clusters to be created
3) creating (centroid/average) of the existing data in each group randomly
4) calculating the distance of each data using the Euclidean distance as follows
5) (Biantara et al., 2023)

$$d = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \quad (1)$$

Note:

d : Euclidean distance

i : number of data points

(x, y): data point

(s, t) : centroid point

6) allocating data to the nearest centroid
7) repeating step 3 if there are still centroids that have changed from the previous centroid
8) Visualizing the cluster results of student scores with criteria of very good and good

d. Cluster Evaluation

The quality of a cluster can be measured by the within-cluster variation, which is the sum of squared errors between all objects and the centroid, defined as follows (Han, J., 2012):

$$E = \sum_{i=1}^{k} \sum_{p \in c_i} dist\ (p, c_i)^2 \quad (2)$$

where E is the sum of squared errors for all objects in the dataset; p is a point in space representing a specific object; and ci is the center of the cluster so that for each object in each cluster, the distance from the object to the cluster center is squared, and these distances are summed. From the formed clusters, the distance can be calculated to determine the objectivity of the dissimilarity between clusters.

FINDINGS

Here are the cluster results of student scores before and after the application of Quizizz

Table 5. Student Score Clusters Before the Application of Quizizz

No	Cluster	Member (Student Attendance Number Order)	Mean	Criteria
1	1	1	94,4	Very Good
2	1	6	94,2	
3	1	7	94	
4	1	8	92,4	
5	1	13	92,2	

6	1	14	90,6	
7	1	17	90,6	
8	1	19	95,2	
9	1	20	92,6	
10	1	21	93,8	
11	2	2	88,6	Good
12	2	3	88,8	
13	2	4	88,4	
14	2	5	88	
15	2	9	89,6	
16	2	10	89,6	
17	2	11	88,4	
18	2	12	88,4	
19	2	15	87	
20	2	16	89	
21	2	18	87,6	
22	2	22	87,2	
23	2	23	89,4	
24	2	24	86,6	

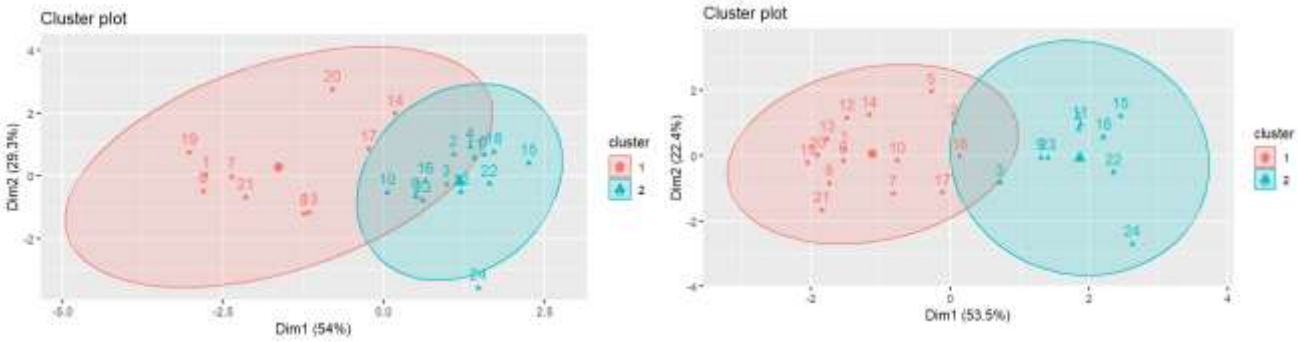
Table 5 shows the score clusters for the English subject before the application of Quizizz, with the average score of students > 90 being 10 students in cluster 1 with a very good criterion, and the average score of students ≤ 90 being 14 students in cluster 2 with a good criterion.

Table 6. Student Score Clusters After the Application of Quizizz

No	Cluster	Member (Student Attendance Number Order)	Mean	Criteria
1	1	1	95,8	
2	1	2	92,6	
3	1	5	93,2	
4	1	6	95,8	
5	1	7	94	
6	1	8	96	
7	1	10	94	
8	1	12	95,8	Very Good
9	1	13	96,4	
10	1	14	95	
11	1	16	92	
12	1	17	92,6	
13	1	19	96,8	
14	1	20	96,6	
15	1	21	96,2	
16	2	3	90	Good

17	2	4	88,4
18	2	9	89,6
19	2	11	88,4
20	2	15	87,2
21	2	18	87,6
22	2	22	87,2
23	2	23	89,4
24	2	24	86,6

Table 6 shows the score clusters for the English subject after the application of Quizizz, with the average score of students > 90 being 14 students in cluster 1 with a very good criterion, and the average score of students ≤ 90 being 10 students in cluster 2 with a good criterion. This indicates an increase of 4 students after the application of Quizizz and an improvement in the average score of other students.



(a) (b)

Figure 1. Classification of Scores a) before and b) after the application of Quizizz in the English subject for students

Based on Figure 1, students in cluster 1 are those with very good average scores in the English subject, and cluster 2 consists of students with good average scores. In Figure 1a, which represents the student cluster before the application of Quizizz, there are 10 students with very good average scores and 14 students with good average scores, while in Figure 1b, which represents the student cluster after the application of Quizizz, the number of students with very good average scores in the English subject increased to 14, and those with good average scores decreased to 10.

Cluster Evaluation

The evaluation can be observed from the average within and average between cluster values. A good cluster is one that has a smaller average within value compared to the average between. Below, Table 7 presents the average within and average between values for K-Means in this study with a total of 2 clusters.

Table 7. Average within and average between values for K-Means

Average within score	Average between score
7.885	13.328

Based on Table 7, it is found that the average within value for K-Means is smaller than the average between value for K-Means, indicating that the cluster model formed is good. A cluster is considered effective or optimal when the data within a cluster are very similar or close to each other (indicating strong internal cohesion) while at the same time, the clusters are far apart from each other (indicating clear differentiation between clusters). This condition shows that the clustering model has successfully identified homogeneous groups of data and significantly differentiated them from other groups of data, which is an indication of good clustering performance.

Discussion

The K-Means method has been extensively utilized in various studies. For instance, the K-Means Cluster Analysis algorithm has been employed in analyzing populations with information technology and communication skills (Abid & Setiawan, 2023). Additionally, K-Means has been used to identify patterns of internet usage among employees at the Research and Development Center for Research and Training of the Indonesian Institute of Sciences (Sukoco, 2018), employing web usage mining with the K-Means clustering algorithm. Furthermore, it has been applied to cluster students based on GPA and several course attributes (Asroni & Adrian, 2016).

This study explores the application of the K-Means algorithm in determining the effectiveness of using Quizizz as a learning medium in teaching English to students at SMP IT Al-Akhyar Makassar. Through the analysis of student learning performance data before and after the application of Quizizz, this research provides how technology, specifically the interactive learning platform Quizizz, can be integrated into the learning process to enhance students' understanding and English language skills. The research findings in this study indicate a significant improvement in student comprehension after the implementation of Quizizz in learning. This is evident from the shift in student allocation to the clusters with "very good" and "good" criteria, where there was an increase in the number of students in the "very good" cluster after the implementation of Quizizz from initially only 10 students to 14 students. Furthermore, as shown in Tables 5 and 6, there was an increase in the average student scores. This indicates that Quizizz positively contributes to the improvement of

students' understanding of English, providing a more engaging and interactive learning environment that supports enhanced learning performance.

Andayani (2007) states that clustering plays a crucial role in various data mining applications, such as knowledge exploration, information retrieval and text mining and web analysis. Clustering is applied in Internet search engines, where search engines will retrieve documents that match the entered keywords. These documents are then grouped into clusters based on the similarity of the words used within them. Cluster analysis is a method that can be utilized to group n number of objects into m clusters. The main objective of cluster analysis is to group several objects based on their similarity in characteristics (Febriani et al., 2021). Moreover, the clustering model evaluation using the K-Means method in this study indicates that the formed clusters have strong cohesion and clear differentiation between clusters, with an average within-cluster distance = 7.885 < average between-cluster distance = 13.328. This confirms that the clustering model successfully identified and grouped student data effectively based on their learning progress, there by proving the effectiveness of using Quizizz in enhancing English language learning.

CONCLUSION

Based on the discussion, it can be concluded that the integration of learning technologies such as Quizizz in English language learning has great potential to enhance students' learning outcomes. This indicates the need for educators to consider using innovative digital learning tools in their teaching strategies to optimize the student learning experience and achieve better learning results. This study also provides insights into the importance of data analysis in education, where the application of data mining methods like K-Means can offer valuable evaluations of the effectiveness of teaching methods and learning tools. Therefore, it is hoped that this research will encourage further exploration of the use of data analysis for educational improvement, as well as the utilization of learning technologies to enrich the English language learning process.

Acknowledgment

The author extends their gratitude to SMP IT Al Akhyar Kota Makassar for providing the opportunity to obtain the data documented in this journal.

REFERENCES

Abid, A., & Setiawan, R. P. (2023). Pemanfaatan Metode Clustering untuk Menganalisis Penduduk Kebumen yang Memiliki Keterampilan Teknologi dan Komunikasi (TIK). *Journal of Data Science Theory and Application*, 02(02), 36–41.

Adianto, S. (2020). Kajian dan Riset Dalam Teknologi Pembelajaran Penerapan Scientific dan Cooperative Learning dengan Quis Online untuk Meningkatkan Hasil. *Jurnal Inovasi Teknologi Pembelajaran*, 7(1), 57–65. https://doi.org/10.17977/um031v7i12020p057

Amalia, D. F. (2020). Quizizz Website as an Online Assessment for English Teaching and Learning: Students' Perspectives. J*o-ELT (Journal of English Language Teaching) Fakultas Pendidikan Bahasa & Seni Prodi Pendidikan Bahasa Inggris IKIP*, 7(1), 1. https://doi.org/10.33394/jo-elt.v7i1.2638

Andayani, S. (2007). Pembentukan cluster dalam Knowledge Discovery in Database dengan Algoritma K-Means. In Seminar Nasional Matematika dan Pendidikan Matematika 2007: Trend Penelitian Matematika dan Pendidikan Matematika di Era Global.

Asroni, A., & Adrian, R. (2016). Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang. *Semesta Teknika*, 18(1), 76–82. https://doi.org/10.18196/st.v18i1.708.

Biantara, B., Rohana, T., & Ratna Juwita, A. (2023). Perbandingan Algoritma K-Means dan DBSCAN untuk Pengelompokan Data Penyebaran Covid-19 Seluruh Kecamatan di Provinsi Jawa Barat. *Scientific Student Journal for Information, Technology and Scienc*e, IV(1), 88–94.

Creswell, J. W. 2014. *Research Design: Qualitative, Quantitative and Mixed Methods Approaches, 4 Edition*. London: Sage

Darma, R. N., & Widiastuty, H. (2023). The Importance Of Learning English At School. *Jurnal Pengabdian Kepada Masyarakat*, 1(1), 23–29. https://doi.org/10.37253/landmark.v1i1.7878

Febriani, Y., Sari, Y. P., & Octaria, D. (2021). Metode K-Means Cluster Untuk Mengelompokkan Kota/Kabupaten di Sumatera Selatan Berdasarkan Produksi Ikan Air Tawar. *Sainmatika: Jurnal Ilmiah Matematika Dan Ilmu Pengetahuan Alam*, 18(2), 175. https://doi.org/10.31851/sainmatika.v18i2.6722

Han, J., Kamber, M., Pei, J. 2012. *Data Mining: Concepts and Techniques*. San Francisco: Morgan Kaufmann Publishers.

Ilyosovna, N. A. (2024). Importance of the English Language in Today's World. *International Journal of English Learning & Teaching Skills*, 6(2), 22–24. https://doi.org/10.15864/ijelts.6205

Khasanah, K., & Lestari, A. (2021). The Effect of Quizizz and Learning Independence on Mathematics Learning Outcomes. Tadris: *Jurnal Keguruan Dan Ilmu Tarbiyah*, 6(1), 63–74. https://doi.org/10.24042/tadris.v6i1.7288.

Magdalena, I., & Mahasiswa 6B PGSD. (2021). *Instrumen Tes dan Non Tes dalam Evaluasi Pembelajaran*. Tangerang: Samudra Biru.

Matondang, Z. (2009). Evaluasi Pembelajaran. Medan: Program Pascasarjana Unimed.

Sugiyono. (2013). *Metode Penelitian Kuantitatif Kualitatif dan R&D*. Bandung: Alfabeta.

Sukoco, S.H. (2018). Analisis Kinerja Pegawai Pusbindiklat Peneliti LIPI Berdasarkan Pola Pemanfaatan Internet Melalui Pendekatan Web Usage Mining [Master's thesis, Institut Pertanian Bogor].

Syamfithriani, T. S., Mirantika, N., & Trisudarmo, R. (2023). Perbandingan Algoritma K-Means dan K-Medoids Untuk Pemetaan Daerah Penanganan Diare Pada Balita di Kabupaten Kuningan. *Jurnal Sistem Informasi Bisnis*, 12(2), 132–139. https://doi.org/10.21456/vol12iss2pp132-139

Zein, S., Sukyadi, D., Hamied, F. A., & Lengkanawati, N. S. (2020). English language education in Indonesia: A review of research (2011-2019). *Language Teaching*, 53(4), 491–523. https://doi.org/10.1017/S0261444820000208

Zhao, F. (2019). Using quizizz to integrate fun multiplayer activity in the accounting classroom. *International Journal of Higher Education*, 8(1), 37–43. https://doi.org/10.5430/ijhe.v8n1p37