Information and Communication Technology Competencies Clustering for students for Vocational High School Students Using K-Means Clustering Algorithm

Muhammad Faisal¹, Nuradin², Fajriana², Zahratul Fitri³

¹Master Student Department of Information Technology, Universitas Malikussaleh, Aceh, Indonesia
²Department of Information Technology, Universitas Malikussaleh, Aceh, Indonesia
³Department of Informatics, Universitas Malikussaleh, Aceh, Indonesia
*Corresponding author E-mail: muhammad.207110201006@mhs.unimal.ac.id

Abstract

The k-Means Clustering algorithm is intended to partition data into one or more groups, where data has similarities in one group and data has differences in another. Information and Communication Technology (ICT) Competency data clustering in educational units is considered necessary to facilitate educational facilitation based on the differences in student abilities, determine advanced ICT guidance groups and become a reference in determining the place of Industrial Work Practices (Prakerin). This study aims to find out how the K-Means Clustering algorithm can be applied in clustering the ICT competencies of students at the State Vocational High School (SMK) 3 Lhokseumawe. The benefits generated in this study are in the form of visualization of data clustering that can help teachers and school management in formulating ICT policies at SMKN 3 Lhokseumawe. The data used in this study is the Information and Communication Technology (ICT) competency test score data for the 2021/2022 academic year. The data was obtained through a competency test process that refers to the Minister of Education and Culture Regulation Number 45 of 2015 concerning the Role of ICT/KKPI Teachers in the Implementation of the 2013 Curriculum where ICT competence includes the skills to search, store, process, present and disseminate data and information. Data processing in this study uses the K-means Clustering method and the RapidMiner application. Data processing using the RapidMiner application starts with data preparation, determining the number of clusters, and configuring the method. This study uses 3 (three) cluster configurations, namely the Very Competent, Competent, and Less Competent clusters. Testing data processing using the RapidMiner application resulted in 80 (eighty) students in cluster_0 with a Very Competent rating, 64 (sixty-four) students in cluster_1 with a Competent rating, and 10 (ten) students in cluster_2 with a Less Competent rating.

Keywords: Data Mining, c, K Means, Information and Communication Technology

1. Introduction

Student profiles related to Information and Communication Technology (ICT) competencies must be mapped by each educational unit. This is due to the development of information technology which has affected all spheres of life, including in the world of education. This condition cannot be denied that every learning and curriculum application process is permanently attached to the Information and Communication Technology (ICT) approach. It is essential to increase the competence of cognitive, affective, and psychomotor learners in solving problems based on information technology [1] [2] [3] [4]. Therefore, educational units that apply a learning approach based on Information and Communication Technology (ICT) can present independent learning innovations compared to conventional learning approaches without collaboration with ICT. Learning through the ICT approach provides a new method in teaching and learning because it can minimize differences in teaching methods and materials, thus providing a more consistent standard of learning quality [5] [6] [7].

The results of previous studies showed that the application of interactive learning with the Blended Learning approach could improve the quality of learning for students at SMKN 3 Lhokseumawe [8]. This makes SMKN 3 Lhokseumawe an educational unit that implements blended learning where the learning process collaborates with ICT, thus making the ICT competence of students an essential aspect to know and map. The results of the ICT competency mapping can be used as a reference by the education unit to formulate lesson plans, formulate ICT policies and become a reference for determining the place for Industrial Work Practices (Prakerin) for students. Therefore, it is deemed necessary to cluster Information and Communication Technology (ICT) competency data at SMKN 3 Lhokseumawe so that
it can automatically generate a visualization of the ICT competency mapping of students. Mapped ICT competencies refer to Permendikbud number 45 of 2015 where ICT competencies are mapped in terms of the ability to search, store, process and disseminate data and information in various ways to support smooth learning. Full online or blended learning is the best solution during the COVID-19 pandemic, causing differences in the learning process. Changes in the learning process are expected to be implemented effectively and efficiently [9].

Data clustering is one of the methods of Data Mining, which is a process that performs one or more computer learning techniques (machine learning) to analyze and extract knowledge automatically. Where the main functions of Data Mining are classification, assessment, prediction, relevance grouping or association rules, grouping, description and visualization [10] [11]. Data classification is seen as very important in the current era of information technology because the Education unit will easily process data and the data security aspect will be guaranteed [12] [13].

The algorithm that can be used to cluster data is the K-Means Clustering method. K-Means Clustering Algorithm is a technique in Data Mining that can partition data into several clusters so that data with the same characteristics will be grouped into one cluster while data with different characteristics will be grouped into other clusters. K-Means performs two processes, namely the process of determining the center of the cluster (centroid) and the process of finding members from each cluster and requires k input parameters or the number of clusters [14] [11] [15].

Previous research using the K-Means Clustering Algorithm was carried out by Miftahul (2021) who examined the implementation of the K-Means Algorithm for Clustering of Participants in the National Science Olympiad at the High School Level which resulted in C1 there were 12 highly competent students to take part in OSN, in C2 there were 14 competent students, but have not been able to take part in OSN and in C3 2 students are less competent to take part in OSN. Ari Sulistiyawati’s research (2021) examines the implementation of the K-Means Clustering Algorithm in the Determination of Superior Class Students, the result is that the implementation of the K-Means Clustering algorithm into the clustering information system provides the results of an effective data grouping classification and the process of each iteration of the centroid distance rotation, point determination clusters are formed, student data as a reference object saves more time on clustering superior classes [16] [17] [15].

Surohman's research (2019) on the use of the K-Means Algorithm in measuring the correlation between student profiles and grades using data on the 2019 National Examination scores at one of the Vocational High Schools in Jakarta obtained the results of the K-Means algorithm that can be used in grouping students based on social aspects. From the research, there were 36% or 184 of the total 512 students whose profile clusters and academic scores matched. 43 students with low grades and minimal family economy and 83 students with the highest academic scores and affluent family economic conditions [18]. Suhefi’s research (2020) on clustering the determination of student interest in school selection uses the K-Means algorithm where in the research the data used is sourced from the Basic Education Data (Dapodik). The result is that the K-Means Clustering algorithm can be used to cluster students’ interest in school selection. From 10 data on schools at the junior high school level, the results obtained for the less desirable category (C0) as many as 6 schools, the moderately attractive category (C1) as many as 3 schools, and the very attractive category (C2) as many as 1 school [19].

This study aims to find the group of Information and Communication Technology (ICT) competency values of SMKN 3 Lhokseumawe students by calculating the shortest distance between the data and the center point (centroid) in a cluster. Thus the intended grouping will produce several clusters, namely Very Competent, Competent and Less Competent in the field of Information and Communication Technology (ICT) [11] [20].

2. Methods

A research requires a framework where the research framework consists of the stages carried out to complete the research. The stages of work in this research can be seen in Figure 1 below:

![Fig 1: Stages of Research Work](image)

Based on the flow of the above work stages can be described as follows:
2.1. Identification of problem

The purpose of this research is to determine the competency clusters of Information and Communication Technology (ICT) students who can determine the location of Industrial Work Practices (Prakerin). Based on the recapitulation of the ICT competency test scores, three clusters were formed, namely Highly Competent, Competent and Less Competent.

2.2. Research Purposes

The purpose of this research is to determine the competency clusters of Information and Communication Technology (ICT) students who can determine the location of Industrial Work Practices (Prakerin). Based on the recapitulation of the ICT competency test scores, three clusters were formed, namely Highly Competent, Competent and Less Competent.

2.3. Data Collecting

The data used in this study is the data of the Information and Communication Technology (ICT) competency test score. Where the data has been encapsulated into variables looking for, storing, processing, presenting and disseminating data or information.

2.4. Data Processing

Data processing in this study uses the K-Means Clustering algorithm, with the following steps:
1. Determine the number of k clusters, where the number of clusters depends on system requirements
2. Determine the centroid value as many as the number of clusters k
3. Allocate each data to the nearest cluster center, the formula used is as follows:

\[ D(x, y) = \sqrt{(x1 - y1)^2 - (xi - yi)^2} \]  

(1)

**Definition 6.1.** Euclidean Distance formula

Where D(x,y) is the data distance to the center of the cluster, x is the data record and y is the data centroid.

4. Recalculate the new centroid, the formula used is as follows:

\[ C1 = \frac{x1 + x2 + x3 + ... + xn}{\sum x} \]  

(2)

Where C1 is the new centroid, x1 is the value of the first data record, xn is the value of the nth data record while x is the number of data records.

5. Repeat process number 4.3. if the cluster center is still changing. If there are no more changes, the data allocation process to the cluster center is stopped.

2.5. Result Testing

Testing the results is intended to obtain accurate data from the system built so that it can run according to research objectives. Data testing is carried out by providing data used when testing applications, determining the results of ICT competency clusters, using the RapidMiner application to build models and visualize data.

2.6. Result Conclusioning

The conclusion is formulated based on the results of the calculation of the K-Means Clustering algorithm using the Rapid Miner application. The results of the clustering of students' ICT competencies will determine where to carry out Industrial Work Practice activities.
3. Results and Discussion

The clustering of Information and Communication Technology (ICT) competencies for SMKN 3 Lhokseumawe students can be explained in the following stages:

1. Recapitulation of Information and Communication Technology (ICT) Competency Test Scores

The amount of data processed in this study was 154 data according to the number of students at level XI at SMKN 3 Lhokseumawe. The ICT competency test scores of these students can be seen in the following table:

**Table 1. Information and Communication Technology (ICT) Competency Test Score Students of SMKN 3 Lhokseumawe – Level XI**

| Name                         | Searching | Storing | Processing | Presenting | Sharing |
|------------------------------|-----------|---------|------------|------------|---------|
| ANJALI DEYA RANI SINAGA      | 88.00     | 79.00   | 71.00      | 76.00      | 63.00   |
| ARIFA MUTIARA                | 86.00     | 80.00   | 81.00      | 78.00      | 78.00   |
| BUNGA SABILA NARIZKI         | 78.00     | 81.00   | 83.00      | 81.00      | 88.00   |
| CUT NURINA FITRI             | 81.00     | 76.00   | 80.00      | 79.00      | 68.00   |
| DINI RIVANI                  | 80.00     | 86.00   | 82.00      | 80.00      | 69.00   |
| EGI VIRLA AKBARI             | 84.00     | 82.00   | 76.00      | 70.00      | 81.00   |
| HAFIZAH JIKRA                | 90.00     | 81.00   | 80.00      | 85.00      | 81.00   |
| INTAN SRI WIDYA              | 90.00     | 79.00   | 76.00      | 87.00      | 83.00   |
| KHANSA ALYANIZAR              | 80.00     | 76.00   | 77.00      | 78.00      | 88.00   |
| LUBNA AL QADRI               | 81.00     | 83.00   | 70.00      | 79.00      | 67.00   |
| MAYATUL HUSNA                | 87.00     | 76.00   | 75.00      | 83.00      | 80.00   |
| MUHAMMAD ALVIAN              | 81.00     | 80.00   | 71.00      | 83.00      | 64.00   |
| MUHAMMAD FITRAH DAVID        | 80.00     | 77.00   | 89.00      | 83.00      | 77.00   |
| NATASYA ZULMAIRA             | 85.00     | 83.00   | 70.00      | 87.00      | 68.00   |
| NURUL IZZA                   | 82.00     | 78.00   | 86.00      | 89.00      | 62.00   |
| PUTRI SALSA BILA             | 83.00     | 86.00   | 70.00      | 85.00      | 84.00   |
| RAIHAN MOULANA               | 81.00     | 76.00   | 76.00      | 84.00      | 64.00   |
| RISKY PRATAMA                | 81.00     | 90.00   | 73.00      | 75.00      | 83.00   |
| ROUZATUL SHA FHITRI          | 90.00     | 78.00   | 83.00      | 80.00      | 82.00   |
| ROZATUL ZANNAH               | 88.00     | 85.00   | 85.00      | 70.00      | 83.00   |
| SITI FAZIA                   | 85.00     | 89.00   | 76.00      | 89.00      | 63.00   |
| SIVA NADILA                  | 90.00     | 87.00   | 79.00      | 81.00      | 89.00   |
| SUCI RAMADHANI               | 88.00     | 76.00   | 76.00      | 81.00      | 86.00   |
| ULYA SALSA BILA              | 85.00     | 79.00   | 79.00      | 77.00      | 65.00   |
| ASMAUL HUSNA                 | 85.00     | 87.00   | 82.00      | 88.00      | 81.00   |
| BINDIA AZZAHRA               | 89.00     | 90.00   | 89.00      | 89.00      | 80.00   |
| ELSA TRIRANDA                | 81.00     | 76.00   | 85.00      | 83.00      | 74.00   |
| FARISYA HUSNA                | 89.00     | 88.00   | 84.00      | 89.00      | 80.00   |
| HETI MAWADDAA                | 90.00     | 87.00   | 88.00      | 84.00      | 81.00   |
| IRFANDI                      | 84.00     | 76.00   | 87.00      | 84.00      | 73.00   |
| KHAILRUL NISA                | 83.00     | 88.00   | 80.00      | 82.00      | 84.00   |
| Name                        | Searching | Storing | Processing | Presenting | Sharing |
|-----------------------------|-----------|---------|------------|------------|---------|
| M. FARHAN MAULANA           | 90.00     | 87.00   | 82.00      | 90.00      | 69.00   |
| MUDA BALIA                  | 83.00     | 84.00   | 90.00      | 84.00      | 79.00   |
| MUHAMMAD ANDIKA             | 85.00     | 80.00   | 82.00      | 76.00      | 76.00   |
| MUHAMMAD HAFIZ              | 81.00     | 85.00   | 78.00      | 78.00      | 78.00   |
| MUHAMMAD HIZAAM             | 83.00     | 84.00   | 82.00      | 87.00      | 66.00   |
| MUHAMMAD HUSAIRI            | 85.00     | 89.00   | 86.00      | 80.00      | 60.00   |
| MUKHTARIZAL                 | 87.00     | 81.00   | 78.00      | 80.00      | 87.00   |
| NITA PERTIWI                | 78.00     | 79.00   | 82.00      | 76.00      | 69.00   |
| PUTRI AZURA                 | 78.00     | 82.00   | 74.00      | 76.00      | 90.00   |
| PUTRI SALSABILA             | 88.00     | 88.00   | 75.00      | 79.00      | 73.00   |
| REVALDY AGUSTRIADI          | 80.00     | 82.00   | 80.00      | 95.00      | 84.00   |
| RISKA NAIWA PUTRI           | 82.00     | 76.00   | 75.00      | 80.00      | 66.00   |
| SITI NUR AZHURA             | 87.00     | 79.00   | 73.00      | 84.00      | 81.00   |
| SRI WAHDINI                 | 81.00     | 79.00   | 87.00      | 85.00      | 82.00   |
| SYAWALUDDIN                 | 87.00     | 75.00   | 87.00      | 78.00      | 60.00   |
| UNC1 AMALIA                 | 80.00     | 82.00   | 80.00      | 80.00      | 80.00   |
| ZUNNUR AINI                 | 86.00     | 88.00   | 80.00      | 80.00      | 70.00   |
| ALDI AGUSTI                 | 79.00     | 86.00   | 90.00      | 90.00      | 82.00   |
| CHAIRUL IKRAMA              | 90.00     | 88.00   | 72.00      | 95.00      | 81.00   |
| DANIL MISWAR                | 81.00     | 88.00   | 90.00      | 79.00      | 61.00   |
| FEBRIANSYAH                 | 90.00     | 95.00   | 88.00      | 91.00      | 92.00   |
| HAYATUL DAMAYANI            | 81.00     | 87.00   | 74.00      | 78.00      | 73.00   |
| ISWANDA                     | 81.00     | 80.00   | 79.00      | 80.00      | 69.00   |
| KAMAL SANDANA               | 87.00     | 88.00   | 90.00      | 78.00      | 86.00   |
| KHAIRIL AKMAL               | 80.00     | 84.00   | 84.00      | 76.00      | 83.00   |
| MUHAMMAD ARIS MAULANA       | 80.00     | 75.00   | 81.00      | 80.00      | 78.00   |
| MUHAMMAD DAFVA              | 80.00     | 85.00   | 88.00      | 81.00      | 89.00   |
| MUHAMMAD HAIQAL             | 80.00     | 79.00   | 75.00      | 84.00      | 60.00   |
| MUHAMMAD MUKSALMINA         | 90.00     | 80.00   | 73.00      | 84.00      | 69.00   |
| MUHAMMAD NAUFAL             | 88.00     | 89.00   | 94.00      | 89.00      | 84.00   |
| MUHAMMAD SYIVA              | 80.00     | 84.00   | 78.00      | 79.00      | 77.00   |
| MUHAMMAD ZIAUDDIN           | 85.00     | 83.00   | 75.00      | 89.00      | 88.00   |
| MUHAMMAD ZULHAM             | 82.00     | 84.00   | 75.00      | 80.00      | 82.00   |
| MUKTARRIZA                  | 81.00     | 76.00   | 88.00      | 86.00      | 82.00   |
| MUNAZILLA RAHMI             | 78.00     | 87.00   | 80.00      | 88.00      | 82.00   |
| MUTIA IBRAHIM               | 83.00     | 82.00   | 86.00      | 84.00      | 64.00   |
| RASHA IMAM MAULANA          | 88.00     | 90.00   | 73.00      | 82.00      | 68.00   |
| RISKA MULFADHILLAH          | 83.00     | 86.00   | 90.00      | 84.00      | 73.00   |
| YASIR ARAFAT                | 83.00     | 79.00   | 90.00      | 84.00      | 88.00   |
| Name                        | Searching | Storing | Processing | Presenting | Sharing |
|-----------------------------|-----------|---------|------------|------------|---------|
| ADAWWIYAH SUFIYANI          | 80.00     | 86.00   | 88.00      | 89.00      | 85.00   |
| ALSYIFA NABILA              | 85.00     | 80.00   | 83.00      | 79.00      | 60.00   |
| ANNISA                      | 85.00     | 86.00   | 70.00      | 89.00      | 79.00   |
| ANNISA AULIA                | 82.00     | 76.00   | 89.00      | 84.00      | 75.00   |
| AVIRATUL AZIMAHE            | 85.00     | 86.00   | 86.00      | 82.00      | 80.00   |
| CUT IZZARA SYAQILLA         | 86.00     | 76.00   | 80.00      | 86.00      | 78.00   |
| DEA DARA PONA               | 88.00     | 76.00   | 73.00      | 89.00      | 83.00   |
| DELVIANI PUTRIA             | 81.00     | 83.00   | 89.00      | 82.00      | 77.00   |
| DIRA SALIMA                 | 81.00     | 81.00   | 78.00      | 77.00      | 85.00   |
| FAHRUL                      | 79.00     | 88.00   | 81.00      | 87.00      | 88.00   |
| FITRI AZ ZAHARA             | 88.00     | 84.00   | 75.00      | 87.00      | 67.00   |
| ISKAL JUA INTAHA            | 79.00     | 84.00   | 87.00      | 89.00      | 90.00   |
| ISMA RAHMATIKA              | 79.00     | 90.00   | 80.00      | 80.00      | 81.00   |
| KHIZAYA AZZURA NASYIFA      | 89.00     | 86.00   | 74.00      | 83.00      | 87.00   |
| MUHAMMAD RAFI QUL A'LA      | 80.00     | 67.00   | 70.00      | 65.00      | 68.00   |
| MUHAMMAD ZIKRI              | 87.00     | 78.00   | 75.00      | 77.00      | 71.00   |
| NABILA SYIFAH               | 90.00     | 79.00   | 76.00      | 83.00      | 74.00   |
| NAHDATUL MUSBAR             | 88.00     | 87.00   | 83.00      | 85.00      | 78.00   |
| NAJWA SYAKBAN NAZILA        | 78.00     | 86.00   | 70.00      | 76.00      | 74.00   |
| NURFAZILA                   | 91.00     | 89.00   | 90.00      | 87.00      | 90.00   |
| NURUL HAYATI                | 80.00     | 90.00   | 82.00      | 79.00      | 76.00   |
| RAHMATUN NISA               | 90.00     | 78.00   | 78.00      | 80.00      | 62.00   |
| RAHMAYANI                   | 87.00     | 82.00   | 74.00      | 77.00      | 71.00   |
| SABNATUL JANNAH             | 87.00     | 86.00   | 78.00      | 82.00      | 79.00   |
| SITI AURA NABILLAH          | 84.00     | 70.00   | 64.00      | 68.00      | 78.00   |
| TEGUH FIRMANSYAH            | 86.00     | 83.00   | 78.00      | 88.00      | 61.00   |
| ZHULHAYYA FITRI             | 90.00     | 78.00   | 80.00      | 77.00      | 70.00   |
| ASSYIFA ANZAR MIATI         | 90.00     | 87.00   | 80.00      | 80.00      | 90.00   |
| CITRA WULANDARI             | 87.00     | 80.00   | 77.00      | 80.00      | 67.00   |
| DHILVI DHANIA               | 90.00     | 78.00   | 80.00      | 90.00      | 82.00   |
| FACHRINA                    | 81.00     | 84.00   | 89.00      | 89.00      | 80.00   |
| FITRAH YANI                 | 90.00     | 77.00   | 77.00      | 78.00      | 73.00   |
| FITRIANI                    | 85.00     | 75.00   | 87.00      | 82.00      | 82.00   |
| HASHIFA NURI                | 86.00     | 75.00   | 90.00      | 83.00      | 68.00   |
| JUNITA MUNASARI             | 89.00     | 84.00   | 81.00      | 84.00      | 65.00   |
| MELISA HAKIKI                | 89.00     | 89.00   | 87.00      | 88.00      | 87.00   |
| MIFTAHUL JANNAH             | 89.00     | 86.00   | 80.00      | 88.00      | 72.00   |
| MUHAMMAD RAIHAN             | 90.00     | 90.00   | 79.00      | 76.00      | 70.00   |
| NURI ASYIDAH                | 83.00     | 80.00   | 84.00      | 75.00      | 63.00   |
| Name                        | Searching | Storing | Processing | Presenting | Sharing |
|-----------------------------|-----------|---------|------------|------------|---------|
| NURUL ALISA                 | 79.00     | 76.00   | 79.00      | 75.00      | 80.00   |
| SALSA NADILA                | 82.00     | 78.00   | 85.00      | 85.00      | 75.00   |
| SEPTIA RAHAYU               | 81.00     | 85.00   | 78.00      | 90.00      | 88.00   |
| SITI FAZIRA                 | 70.00     | 65.00   | 74.00      | 65.00      | 75.00   |
| ZUHRA ULVINA                | 84.00     | 87.00   | 72.00      | 75.00      | 66.00   |
| ZULFA SALMA FATIN           | 83.00     | 78.00   | 71.00      | 77.00      | 79.00   |
| ALDY SYALWANI               | 86.00     | 86.00   | 87.00      | 86.00      | 67.00   |
| ARIEL SYAHREZA              | 88.00     | 83.00   | 84.00      | 85.00      | 67.00   |
| BASYIRUN NAZIR              | 83.00     | 83.00   | 77.00      | 78.00      | 65.00   |
| FAKRUL RAZI                 | 84.00     | 87.00   | 90.00      | 79.00      | 79.00   |
| FATURROZI                   | 87.00     | 78.00   | 81.00      | 90.00      | 85.00   |
| FIRMAN ALFATHY              | 80.00     | 80.00   | 78.00      | 88.00      | 82.00   |
| MARSIDA ARISTA              | 89.00     | 87.00   | 72.00      | 84.00      | 72.00   |
| MUHAMMAD DAFFA ANNAS        | 85.00     | 88.00   | 77.00      | 89.00      | 74.00   |
| MUHAMMAD DANIL              | 89.00     | 88.00   | 75.00      | 88.00      | 86.00   |
| MUHAMMAD FAIZIR GIBRAN      | 89.00     | 83.00   | 74.00      | 86.00      | 87.00   |
| MUHAMMAD FANNY              | 90.00     | 88.00   | 80.00      | 84.00      | 60.00   |
| MUHAMMAD HABIL              | 81.00     | 76.00   | 85.00      | 90.00      | 73.00   |
| MUHAMMAD HAIFIDZ            | 81.00     | 80.00   | 75.00      | 76.00      | 72.00   |
| MUHAMMAD HARDIANSYAH        | 83.00     | 79.00   | 82.00      | 82.00      | 69.00   |
| MUHAMMAD IKBAL              | 79.00     | 76.00   | 70.00      | 75.00      | 87.00   |
| MUHAMMAD IKHFAR             | 86.00     | 81.00   | 72.00      | 77.00      | 61.00   |
| MUHAMMAD KHADAFI            | 79.00     | 79.00   | 87.00      | 85.00      | 63.00   |
| MUHAMMAD RAFLI              | 79.00     | 89.00   | 87.00      | 85.00      | 71.00   |
| MUHAMMAD RAIYAN AKBAR       | 83.00     | 90.00   | 86.00      | 78.00      | 70.00   |
| MUHAMMAD REHAN              | 79.00     | 83.00   | 75.00      | 77.00      | 87.00   |
| MUHAMMAD RIFAT ALFIZAL      | 84.00     | 85.00   | 83.00      | 76.00      | 74.00   |
| MUHAMMAD RISKI              | 84.00     | 86.00   | 79.00      | 85.00      | 60.00   |
| RAHMAWATI                   | 90.00     | 78.00   | 85.00      | 83.00      | 80.00   |
| REVA FAZILLAH IZZATI        | 81.00     | 83.00   | 73.00      | 89.00      | 75.00   |
| SYIFA ALIYA                 | 82.00     | 80.00   | 88.00      | 83.00      | 69.00   |
| TURHAMUN                    | 90.00     | 85.00   | 87.00      | 86.00      | 76.00   |
| ZAHWA                       | 83.00     | 88.00   | 77.00      | 86.00      | 63.00   |
| ZAWIRA ANNISA               | 79.00     | 78.00   | 73.00      | 78.00      | 76.00   |
| ZULFIKAR                    | 86.00     | 77.00   | 90.00      | 82.00      | 69.00   |
| AURA MAULIZA                | 87.00     | 76.00   | 79.00      | 78.00      | 80.00   |
| AUREL RAYA RAMADHANI        | 87.00     | 80.00   | 80.00      | 85.00      | 87.00   |
| FADILA NUFUS                | 88.00     | 86.00   | 82.00      | 75.00      | 87.00   |
| FIKI MAULANDA               | 87.00     | 85.00   | 81.00      | 84.00      | 76.00   |
2. Determine the number of clusters
The number of clusters in this study amounted to three clusters, namely Very Competent, Competent and Less Competent.

3. Calculating the data distance with the centroid using the Rapid Miner application
Calculating the data distance with the centroid using the Rapid Miner application, RapidMiner is a reliable software or application in data processing. Based on data mining principles and algorithms, RapidMiner extracts patterns from large data sets and combines statistical methods, artificial intelligence and databases [21]. RapidMiner can be a solution for analyzing data mining and predictive analysis [22]. Data processing of Information and Communication Technology (ICT) competencies for SMKN 3 Lhokseumawe students uses data and operator configurations, as shown in the following figure:

![Fig 2: Data and operator configuration](image)

The Information and Communication Technology (ICT) competency test score file is imported into the RapidMiner application worksheet then the configuration is followed by importing the K-Means algorithm operator and Cluster Distance Performance segmentation. In the K-Means algorithm operator, 3 (three) clusters are determined and in the performance section the main criterion selected is Davies Bolubidin. Once the operator configuration and performance are confirmed, the process can be started. The results of the data processing are displayed in the form of Cluster Model, Clustering Example Set and Performance Vector. In the Cluster Model section, you can see the results of the cluster names of participants who have been divided into 3 clusters and the cluster data distribution plot, as shown below:

![Fig 3: Folder View pada Cluster Model RapidMiner](image)

Based on the results of the data processing using 3 (three) clusters, it is obtained that cluster_0 contains 80 students, cluster_1 contains 64 students and cluster_2 contains 10 students. The results of the process through RapidMiner can be seen in the following image:
Based on the resulting cluster model, only the number of members in the cluster can be seen. To find out which cluster has the highest score group, the plot features in the cluster model can be used. The data plot can be seen in the following figure:

![Cluster Model Process Results](image1)

**Fig 4:** Cluster Model Process Results

Through the picture above, it can be seen that cluster_0 is identified using blue where the cluster is at the highest level, meaning that cluster_0 contains data on students who have a high level of competence (very competent). Cluster_1 is an intermediate (competent) cluster identified in green while cluster_2 is identified in red as a cluster of students who have a low level of ICT competence (less competent).

4. Conclusion

Based on the results of the data processing of the Information and Communication Technology (ICT) competency test scores 154 students of SMKN 3 Lhokseumawe using the RapidMiner application with a configuration of 3 (three) clusters resulting in 80 (eighty) students in cluster_0 with a rating of Very Competent, 64 (sixty four) students in cluster_1 with a Competent rating and 10 (ten) students in cluster_2 with a Less Competent rating. Based on the results of the mapping, the education unit can formulate a policy for Industrial Work Practices (Prakerin) for students in each cluster.

References

[1] Z. Fitri and E. Utaminingsih, “Penerapan Metode Computational Thinking Pada Kurikulum Aceh Untuk Mencapai Kognitif,” *J. MathEducation Nusant.*, vol. 4, no. 1, pp. 60–73, 2021.

[2] Yusrizal, “SURVEY PENGGUNAAN TEKNOLOGI INFORMASI DAN KOMUNIKASI DI WILAYAH PERBATASAN SURVEY,” *J. Teknol. Inf. dan Komun. Inf. dan Komun.*, 2016.

[3] H. Chung, “Sikap guru Bahasa Melayu terhadap penggunaan teknologi maklumat dan kKomunikasi (ICT) dalam pengajaran di Sekolah-sekolah Rendah di Bintulu, Sarawak,” *J. Pendidik. Malaysia*, 2010.

[4] R. Siva Sankara Raju, A. D, and P. Thimothy, “Determination of Stress and Deformations Analysis on LPG Steel Cylinder,” *Int. J. Eng. Res. Appl.*, vol. 3, no. 1, 2013.

[5] Fajriana, “Pemanfaatan Teknologi Informasi Untuk Meningkatkan Kualitas Pembelajaran,” *TECHSI*, vol. 9, 2017.
Z. Vahedi, L. Zannella, and S. C. Want, “Students’ use of information and communication technologies in the classroom: Uses, restriction, and integration,” Act. Learn. High. Educ., vol. 22, no. 3, 2021, doi: 10.1177/1469787419861926.

W. M. Al-Rahmi, A. I. Alzahrani, N. Yahaya, N. Alalwan, and Y. Bin Kamin, “Digital communication: Information and communication technology (ICT) usage for education sustainability,” Sustain., vol. 12, no. 12, 2020, doi: 10.3390/su12125052.

Z. Fitri, U. A., and M. Z. Akbar, “Penerapan Media Pembelajaran Interaktif Berbasis Blended Untuk Meningkatkan Kualitas Belajar Siswa Di Smkn 3 Lhokseumawe,” J. Sist. Inf. ISSN, vol. 1, no. Interaktif Learning, pp. 1–6, 2018.

Nurdin, N. Putri, M. Uci, Al-Kautsar, H. Aidilof, and Bustami, “Implementation of Fuzzy C-Means to Determine Student Satisfaction Levels in Online Learning,” Sistemasi, vol. 11, no. 1, p. 121, 2022, doi: 10.32520/stmsi.v11i1.1638.

Nurdin and D. Astika, “Penerapan Data Mining Untuk Menganalisis Penjualan Barang dengan Menggunakan Metode Apriori pada Supermarket Sejahtera Lhoksumawe,” TECHSI, vol. 6, pp. 77–80, 2015.

R. Hutagalung and A. Gustomo, “Workload Analysis for Planning Needs of Employees in The Corporate Administration Unit PT Timah (Persero) Tbk.,” Indones. J. Bus. Adm., vol. 2, no. 19, pp. 2290–2297, 2013.

Nurdin, M. Suhendri, Y. Afrilia, and Rizal, “Klasifikasi Karya Ilmiah (Tugas Akhir) Mahasiswa Menggunakan Metode Naive Bayes Classifier (NBC),” Sistemasi, vol. 10, no. 2, p. 268, 2021, doi: 10.32520/stmsi.v10i2.1193.

S. Surohman, L. Fabrianto, F. Riza, and N. M. Faizah, “Korelasi Antar Profil dan Nilai Akademis Siswa dengan Menggunakan Algoritma K-Means,” J. Teknol. dan Sist. Inf., vol. IV, no. 1, 2017.