

Penerapan Tingkat Minat Baca Siswa SD Inpres 037145 Menggunakan *K-Means***Anita¹, Herwin Simbolon²**Program Studi Sistem Informasi Fakultas Teknologi dan Ilmu Komputer Universitas Prima
Indonesia^{1,2}Email : anitayakub_pilchan@yahoo.com¹DOI: <http://dx.doi.org/10.31869/rtj.v4i2.2556>

Abstract: Grouping students' reading interest based on the criteria for books read, books borrowed, and also considering the number of books available can help in the process of adding to the existing book collection in the library. One way to manage this data is by using data mining by utilizing the *K-Means* method. Book data are grouped into 3 clusters, namely priority, consideration, and not prioritization in planning for additional book collections. The result of this research is that the cluster with the largest value in the final centroid is the recommended cluster in planning to add to the book collection.

Keywords: clustering, k-means, data, library

Abstrak: Pengelompokan minat baca mahasiswa berdasarkan kriteria buku yang dibaca, buku yang dipinjam, dan juga mempertimbangkan jumlah stok buku yang tersedia dapat membantu dalam proses penambahan koleksi buku yang telah ada pada perpustakaan. Salah satu cara untuk mengelola data tersebut yaitu menggunakan data mining dengan memanfaatkan metode *K-Means*. Data buku dikelompokkan menjadi 3 cluster yaitu prioritas, dipertimbangkan, dan tidak diprioritaskan dalam perencanaan penambahan koleksi buku. Hasil dari penelitian ini adalah cluster dengan nilai terbesar pada centroid akhir merupakan cluster yang direkomendasikan dalam perencanaan penambahan koleksi buku.

Kata Kunci: clustering, k-means, data, perpustakaan

PENDAHULUAN

Minat baca merupakan suatu pengaruh yang besar bagi siswa. Dalam meningkatkan upaya minat baca siswa sekolah dasar menjadi tanggung jawab bersama, antara siswa itu sendiri, guru maupun orang tua. Namun rendahnya minat baca siswa di sekolah dasar menjadi salah satu halangan, kurangnya keinginan dan kemauan dari siswa itu sendiri, guru juga belum mengharuskan siswa untuk membaca buku saat pembelajarn yang diajarkan kepada siswa. Dalam upaya meningkatkan minat baca siswa sekolah dasar sebaiknya siswa diberi dukungan agar minat baca itu muncul dari diri siswa dan perpustakaan sebagai tempat menyimpan koleksi bahan pustaka sumber informasi.

Pada Saat ini, sistem yang diterapkan di perpustakaan SD INPRES 037145 berjalan kurang efektif. Ini terlihat dari mulai tahap perencanaan yang meliputi pemilihan hingga penambahan koleksi buku yang tidak maksimal sehingga minat baca siswa berkurang. Salah satu cara untuk mengelola data tersebut yaitu dengan menggunakan data mining. Penerapan data mining telah banyak digunakan dalam pengolahan data untuk

menghasilkan pengetahuan, salah satunya [1] Algoritma *K-Means Clustering* yang telah digunakan dalam berbagai kasus pengelompokan data untuk peminatan literasi [2], Selain itu untuk mengetahui proporsi mengelompokkan literasi yang layak dalam SD INPRES 037145 [3] sehingga *k-means* dapat sebagai alat bantu untuk menentukan literasi yang bagus.

METODOLOGI PENELITIAN**Data Mining**

Data Mining merupakan suatu istilah yang digunakan untuk mendapatkan pengetahuan yang tersembunyi dari kumpulan data yang berukuran sangat besar. yang tujuan utama *Data Mining* adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki[4].

Metode *K-Means Clustering*

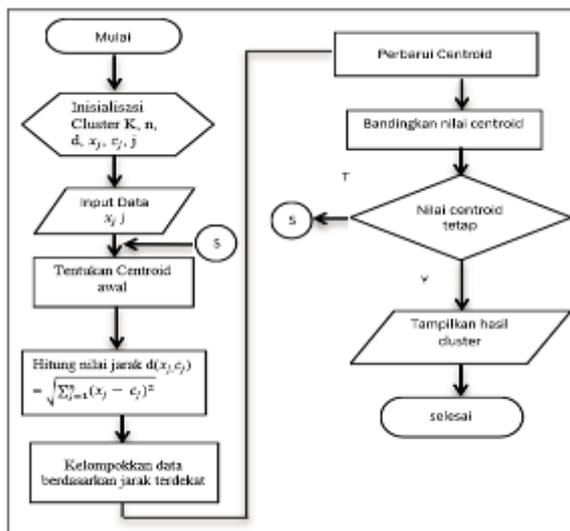
Metode *K-means clustering* merupakan metode *clustering* yang dikenalkan oleh Metode *K-means* adalah metode yang terkenal cepat dan simpel. *K-means clustering* merupakan salah satu metode data *clustering*

non hirarki yang mengelompokan data dalam bentuk satu atau lebih *cluster*/kelompok. Data-data yang memiliki karakteristik yang sama dikelompokan dalam satu *cluster*/kelompok dan data yang memiliki karakteristik yang berbeda dikelompokan dengan *cluster*/kelompok yang lain sehingga data yang berada dalam satu *cluster*/kelompok memiliki tingkat variasi yang kecil [5]. Langkah-langkah melakukan *clustering* dengan metode *K-means* adalah sebagai berikut:

1. Pilih jumlah *cluster* *k*.
2. Inisialisasi ke pusat *cluster* ini bisa dilakukan dengan berbagai cara. Cara yang paling sering dilakukan adalah dengan *random* atau acak. Pusat-pusat *cluster* diberi dengan nilai awal dengan angka-angka *random*
3. Alokasikan semua data/objek ke *cluster* terdekat. Kedekatan dua objek ditentukan berdasarkan jarak ke-dua objek tersebut.

HASIL DAN PEMBAHASAN

Dari langkah-langkah dan urutan prosedur dari suatu program kerja secara keseluruhan menggunakan metode *K-Means* mulai dari awal sampai akhir prosesnya. Berikut ini adalah gambar Flowchartnya algoritma *K-Means* yaitu sebagai berikut :



Gambar 1 Flowchart Algoritma *K-Means*

Data yang menyangkut dengan Peminatan literasi baca SD INPRES 037145 adalah sebagai berikut :

Tabel 1 Data Laporan Jumlah Siswa Peminatan Literasi baca Buku Dalam Peminjaman

No	Nama Siswa	Buku Mata Pelajaran	Buku Cerita	Buku Umum
1	M Zulfikar Surya	1	1	1
2	Mhd Khairuddin Lubis	4	3	2
3	Rusdina Hatimah	10	7	8
4	Desi Wrdani Samara	2	4	6
5	Anto Syahputra Barus	3	4	3
6	Muhammad Rizki Fajar Finky	1	1	1
7	Handayani	2	1	5
8	Fahry Muhammad	1	1	1
9	Alfredo M Sembiring	2	5	5
10	Agustina Ratna Sari	3	5	5
Lanjutan sampai nomor 150				
150	Dede Amelia	2	4	6

Data pada tabel data 1 di atas agar dapat diolah menggunakan algoritma *K-Means*, maka dinormalisasikan atau membersihkan data yang tidak digunakan serta menginisialkan data pengiriman dan alamat dengan ketentuan, Berikut ini langkah-langkah pada algoritma *K-Means* sampai diketahui pembagian nilai *Centroid* sebelumnya tidak berubah.

1. Menentukan jumlah *Cluster* misalkan sebanyak *k* = 3
2. Menentukan *Centroid* *c* setiap *Clusteryang* diambil dari data sumber.

Table 2 Tabel Data *Centroid* Awal

<i>Centroid</i>	Pembelajaran Umum	Buku Cerita	Fiksi
<i>Centroid</i> 1	1	1	1
<i>Centroid</i> 2	3	4	3
<i>Centroid</i> 3	10	7	8

Hitung jarak data ke *Centroid* menggunakan rumus *Euclidean*, data tersebut dari *Cluster* terdekatnya.

a. Jarak antara siswa nomor pertama dengan titik m1

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(1 - 1)^2 + (1 - 121)^2 + (1 - 21)^2}$$

$$= 0$$

b. Jarak antara siswa nomor kedua dengan titik m1

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(4-1)^2 + (3-1)^2 + (2-1)^2}$$

$$= 4,123$$

c. Jarak antara siswa nomor ketiga dengan titik m1

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(10-1)^2 + (7-1)^2 + (8-1)^2}$$

$$= 12,884$$

d. Jarak antara siswa nomor pertama dengan titik m2

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(1-3)^2 + (1-4)^2 + (1-3)^2}$$

$$= 3,742$$

e. Jarak antara siswa nomor kedua dengan titik m2

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(4-3)^2 + (3-4)^2 + (2-3)^2}$$

$$= 1,732$$

f. Jarak antara siswa nomor ketiga dengan titik m2

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(10-3)^2 + (7-4)^2 + (8-3)^2}$$

$$= 9,381$$

g. Jarak antara siswa nomor pertama dengan titik m3

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(1-10)^2 + (1-7)^2 + (1-8)^2}$$

$$= 12,884$$

h. Jarak antara siswa nomor kedua dengan titik m3

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(4-10)^2 + (3-7)^2 + (2-8)^2}$$

$$= 9,110$$

i. Jarak antara siswa nomor ketiga dengan titik m3

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(10-10)^2 + (7-7)^2 + (8-8)^2}$$

$$= 0$$

Untuk lebih lengkapnya jarak pada setiap baris data, hanya ditampilkan 10 dari 150 data seperti pada tabel berikut:

Tabel 3 Tabel Hasil Perhitungan Iterasi Ke 1 (Lanjutan)

No	Nama Siswa	C1	C2	C3	JARAK TERDEKAT
1	M Zulfikar Surya	0,000	4,123	12,884	C1
2	Mhd Khoiruddin Lubis	3,742	1,732	9,381	C2
3	Rusdina Hatimah	12,884	9,110	0,000	C3
4	Desi Widani Samara	5,916	3,162	8,775	C2
5	Anto Syahputra Baras	4,123	0,000	9,110	C2
6	Muhammad Rizki Fajar	0,000	4,123	12,884	C1
7	Fauky Handayani	4,123	3,742	10,440	C2
8	Fahry Muhammad	0,000	4,123	12,884	C1
9	Alfredo M Sembiring	5,745	2,449	8,775	C2
10	Agustina Rama Sari	6,000	2,236	7,874	C2
Lanjutan sampai nomor 150					
150	Dede Amelia	5,916	3,162	8,775	C2

Dari tabel 3 di dapat Jumlah Siswa sebagai berikut :

- C1 = {
1,6,8,13,15,18,22,25,27,31,32,39,
47,48,55,63,64,71,79,80,87,95,96,
103,111,112,119,126,127,134,138,
139,146}
- C2 = {
2,4,5,7,9,10,11,12,14,16,17,19,20,
21,23,24,26,28,29,30,33,34,35,36,
37,38,40,41,43,44,45,46,49,50,51,
53,54,56,57,59,60,61,62,65,66,67,
69,70,72,73,75,76,77,78,81,82,83,
85,86,88,89,91,92,93,94,97,98,99,
101,102,104,105,107,108,109,110,
113,114,115,117,118,120,121,12
3,124,125,128,129,130,132,133,1
35,136,137,140,141,142,144,145,
147,148,150}
- C3 = {
3,42,52,58,68,74,84,90,100,106,1
16,122,131,143,149}
- Lakukan pembaruan *Centroid* dari hasil *Cluster* seperti berikut :
- C1 = rata-rata
(1,6,8,13,15,18,22,25,27,31,32,39,

- 47,48,55,63,64,71,79,80,87,95,96,
103,111,112,119,126,127,134,138
,139,146)
= (1;1;1)
- C2 = rata-rata
(2,4,5,7,9,10,11,12,14,16,17,19,20,2
1,23,24,26,28,29,30,33,34,35,36,37,
38,40,41,43,44,45,46,49,50,51,53,5
4,56,57,59,60,61,62,65,66,67,69,70,
72,73,75,76,77,78,81,82,83,85,86,8
8,89,91,92,93,94,97,98,99,101,102,
104,105,107,108,109,110,113,114,1
15,117,118,120,121,123,124,125,12
8,129,130,132,133,135,136,137,140
,141,142,144,145,147,148,150)
= (2.66;4.00;4.55)
- C3 =rata-rata
(3,42,52,58,68,74,84,90,100,106,1
16,122,131,143,149)
= (10;7;8)

Menghitung kembali nilai rasio dengan membandingkan nilai *BCV* dan *WCV*.

$$BCV/WCV = 26,118 / 225,602 = 0,116$$

Nilai *Centroid* berubah dari nilai *Centroid* sebelumnya, maka algoritma dilanjutkan ke langkah berikutnya.

Hasil pengelompokan *cluster* dari minat baca literasi baca rendah adalah sebagai berikut:

Tabel 4 Pengelompokan Hasil Cluster

Cluster	Nomor
Siswa Peminatan Literasi baca Rendah	1,6,8,13,15,18,22,25,27,31,32,39,47,48,55,63,64,71,79,80,87,95,96,103,111,112,119,126,127,134,138,139,146
Siswa Peminatan Literasi baca Sedang	2,4,5,7,9,10,11,12,14,16,17,19,20,21,23,24,26,28,29,30,33,34,35,36,37,38,40,41,43,44,45,46,49,50,51,53,54,56,57,59,60,61,62,65,66,67,69,70,72,73,75,76,77,78,81,82,83,85,86,88,89,91,92,93,94,97,98,99,101,102,104,105,107,108,109,110,113,114,115,117,118,120,121,123,124,125,128,129,130,132,133,135,136,137,140,141,142,144,145,147,148,150
Siswa Peminatan Literasi baca Tinggi	3,42,52,58,68,74,84,90,100,106,116,122,131,143,149

Tampilan Program

1. *Form Login*

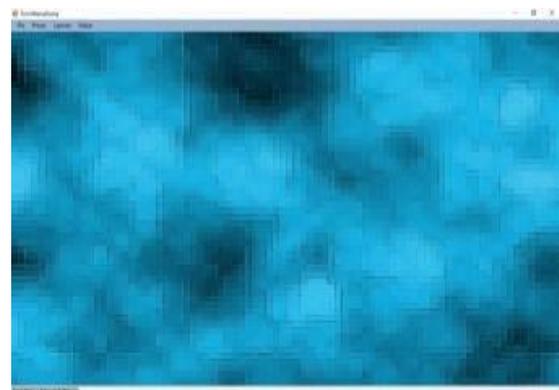
Form Login merupakan *form* untuk melakukan pengisian data awal *user* sebelum masuk ke *Form Menu Utama*. *Form Login* ini bertujuan untuk mengamankan sistem dari *user-user* yang tidak bertanggung jawab. Pada *form* ini, yang dilakukan adalah menginput *User* dan *Password* dengan benar dan sesuai dengan data yang telah di daftarkan pada database *login*. Berikut adalah tampilan *Form Login* :



Gambar 2 Login

2. *Form Menu Utama*

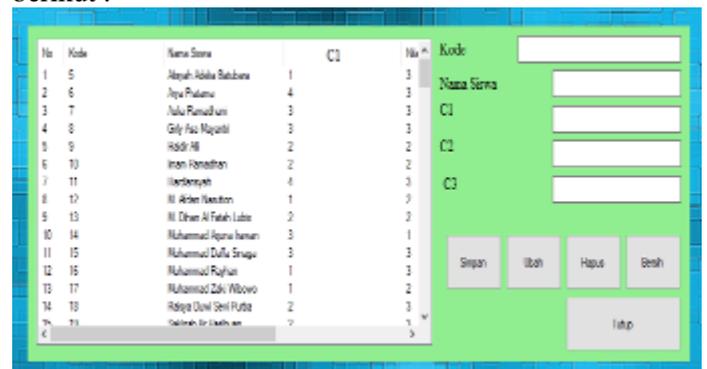
Form Menu Utama digunakan sebagai penghubung untuk *Form Data Siswa*, *Form Input Data Centroid*, *Form* proses *K-Means*, dan *Form Laporan Cluster*. Di Dalam *Form Menu Utama* ini terdapat beberapa menu yang dimana diantaranya adalah menu Data Untuk menampilkan *Form Data Siswa* dan *Form Input Data Centroid*, menu proses untuk menampilkan *Form Proses K-Means*, menu *Laporan Cluster* untuk menampilkan *Form Laporan Hasil Perhitungan Proses K-Means* dan Menu *Keluar* untuk menutup aplikasi. Berikut adalah tampilan *Form Menu Utama* :



Gambar 2 Tampilan Utama

3. *Form Data Siswa*

Form Data Siswa adalah *Form* pengolahan Data Siswa dalam menginput data, menyimpan data, mengubah data dan menghapus data. Adapun *Form Data Siswa* adalah sebagai berikut :



Gambar 4 Tampilan Data Siswa

4. Form Input Nilai Centroid

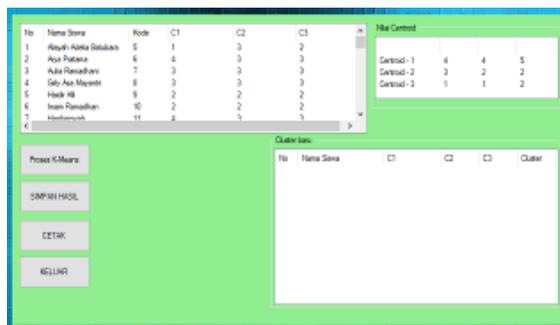
Form Input Nilai Centroid adalah pengolahan data Centroid yang dapat di ubah jika Nilai Centroidnya ditentukan dengan Nilai Centroid yang berbeda. Adapun Form Input Nilai Centroid adalah sebagai berikut :



Gambar 5 Tampilan Input Nilai Centroid

4. Form Metode K-Means

Form Metode K-Means adalah proses perhitungan dalam mengelompokkan data dalam mengelompokkan data berdasarkan variabel yang sudah ditentukan. Adapun Form Metode K-Means adalah sebagai berikut :



Gambar 6 Tampilan Metode K-Means

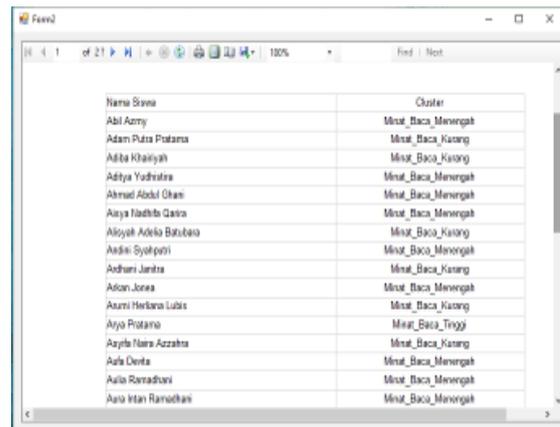
5. Form Cluster

Form Cluster adalah proses pengelompokkan data cluster sesuai variabel yang sudah ditentukan. Adapun Form Cluster adalah sebagai berikut :

Row No.	id	cluster	C1	C2	C3
104	104	cluster_0	3	1	2
105	105	cluster_1	3	2	2
106	106	cluster_2	2	3	2
107	107	cluster_1	4	3	4
108	108	cluster_2	2	3	2
109	109	cluster_0	2	2	2
110	110	cluster_0	3	2	1
111	111	cluster_0	3	1	3
112	112	cluster_0	1	2	3
113	113	cluster_2	1	3	2
114	114	cluster_1	3	3	3
115	115	cluster_2	2	3	3
116	116	cluster_1	4	2	2
117	117	cluster_0	1	2	2
118	118	cluster_0	2	2	2
119	119	cluster_1	3	2	2
120	120	cluster_1	3	3	4

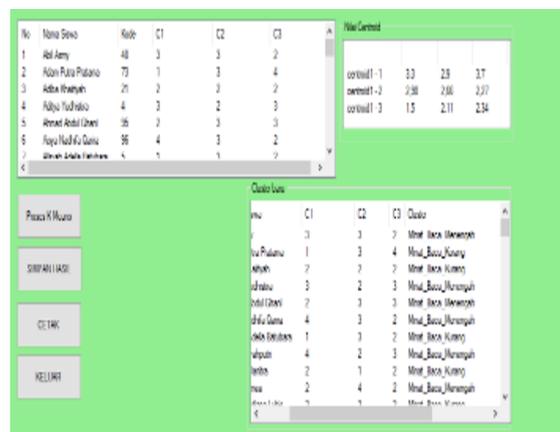
Gambar 7 Tampilan Cluster

6. Data Cluster



Gambar 8 Tampilan hasil Cluster

7. Hasil yang ditampilkan di program sistem Berikut ini adalah hasil dari tampilan di program sistem, adalah sebagai berikut :



Gambar 9 Tampilan hasil yang ditampilkan di program sistem

PENUTUP
Simpulan

Berdasarkan hasil analisa dari permasalahan yang terjadi dengan kasus yang di bahas tentang mengelompokkan minat baca dengan menerapkan algoritma k-means terhadap sistem yang dirancang dan dibangun maka dapat ditarik kesimpulan sebagai berikut:

1. Dalam mengelompokkan metode k-means dapat digunakan untuk mempersingkat siswa mana yang berminat baca dan mengetahui buku apa yang diminati oleh siswa.
2. Dapat merancang sistem dengan menggunakan Class diagram, alur sistem dengan aktor dengan menggunakan activity diagram ataupun use case

diagram dan alur algoritma sistem menggunakan *flowchart*.

3. Dapat menguji sistem dengan menggunakan bahasa pemrograman *visual basic* dan mempercepat dalam pengelompokan data.

Saran

Untuk meningkatkan kemampuan dan fungsi dari sistem ada beberapa saran yang dapat diberikan untuk pengembangan yang bisa dilakukan yaitu :

1. Sistem yang dirancang dan dibangun harus dikembangkan lagi dengan berbasis *Mobile* dan *web*.
2. Disarankan sistem tidak hanya menggunakan algoritma k-means akan tetapi bisa dipadukan dengan algoritma yang lain ataupun dengan kombinasi yang lain untuk meningkatkan keakuratan dalam pengelompokan.
3. Disarankan data yang digunakan dengan menggunakan lebih dari 1 tempat riset yang dikelompokkan dapat meningkatkan tingkat keakuratan pengelompokan.

DAFTAR PUSTAKA

- [1] D. Sunia, K. and A. P. Jusia, "Penerapan *Data mining* Untuk Clustering Data Penduduk Miskin Menggunakan Algoritma K-Means," *STIKOM Dinamika Bangsa*, pp. 121-134, 2019.
- [2] S. S. Helma, M. R. R. R and E. Normala, "Clustering pada Data Fasilitas Pelayanan Kesehatan Kota Pekanbaru Menggunakan Algoritma K-Means," *Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI)*, vol. I, no. 1, pp. 131-137, 2019.
- [3] F. Yunita, "Penerapan *Data mining* Menggunakan Algoritma K-Means Clustering Pada Penerimaan Mahasiswa Baru (Studi Kasus : Universitas Islam Indragiri)," *Jurnal SISTEMASI*, pp. 238-249, 2018.
- [4] Alfannisa Annurullah Fajrin and Algifanri Maulana, "Penerapan *Data mining* Untuk Analisis Pola Pembelian Konsumen Dengan Algoritma Fpgrowth Pada Data Transaksi Penjualan Spare Part Motor," *Kumpulan jurnal Ilmu Komputer (KLIK)*, vol. Volume 05, 2018
- [5] Joko Suntoro, *Data mining: Algoritma Dan Implementasi Dengan Pemrograman*, Elex Media Komputindo, 2019, p. 192.
- [6] S. Kasiyun, "Upaya Meningkatkan Minat Baca Sebagai Sarana Untuk Mencerdaskan Bangsa," *JURNAL PENA INDONESIA (JPI)*, vol.I No.1, 2015
- [7] A. Kurniawan , D. , S. Hayati, R. , J. Riskayanti, I. Wasena, Y.Triadi., "Peranan Pojok Baca dalam Menumbuhkan Minat Baca Siswa Sekolah Dasar," *JURNAL INOVASI PENDIDIKAN DAN PEMBELAJARAN SEKOLAH DASAR*, vol.III No.2, 2019