

**PREDIKSI NILAI UNAS SISWA SMP MENGGUNAKAN METODE
K – MEANS CLUSTERING DAN REGRESI
LINIER SEDERHANA**

¹Yogi Prasetyo (1010651112) ²Daryanto (NPK.1103589) ³Lutfi Ali Muharom (NPK.1009550)

Jurusan Teknik Informatika Fakultas Teknik

Universitas Muhammadiyah Jember

Email : yogi.henshin@gmail.com

ABSTRAK

Sekolah Menengah Pertama (SMP) merupakan sekolah wajib belajar bagi setiap warga negara berusia 7 – 15 tahun di Indonesia. Oleh karena itu perlu untuk meningkatkan *kualitas* dan *kuantitas* Sekolah Menengah Pertama (SMP) yaitu dengan mengembangkan suatu aplikasi yang dapat di gunakan untuk memprediksi nilai UNAS siswa Sekolah Menengah Pertama (SMP) sebagai bahan evaluasi guru untuk menghadapi UNAS siswa Sekolah Menengah Pertama (SMP). Sebagai inputan digunakan nilai ulangan harian dan nilai UTS dari 4 mata pelajaran yaitu Bahasa Indonesia, Bahasa Inggris, Matematika dan IPA yang akan di proses yaitu dengan menggunakan metode *K-Means Clustering* untuk proses pengelompokkan data. Selanjutnya digunakan metode *Regresi Linier Sederhana* untuk mendapatkan nilai prediksi nilai UNAS siswa dengan inputan hasil cluster dari proses *K-Means Clustering* dan nilai UAS siswa. Hasil dari nilai prediksi tersebut kemudian di proses kembali menggunakan *RMSE (Root Mean Squared Error)* untuk mengetahui tingkat *akurasi* prediksi nilai UNAS siswa. Setelah menggunakan metode K-Means dengan beberapa centroid dan *Regresi Linier Sederhana* didapatkan hasil prosentase tingkat akurasi prediksi nilai UNAS siswa yang terkecil dengan metode RMSE (Root Means Squared Error) pada mata pelajaran bahasa Indonesia adalah 41.5, bahasa Inggris adalah 51.3, IPA adalah 13.8, dan Matematika adalah 44.3 .

Kata Kunci : *SMP, UNAS, K-Means Clustering, Regresi Linier Sederhana, RMSE.*

PENDAHULUAN

1.1 Latar Belakang

Pendidikan merupakan salah satu elemen terpenting suatu negara, tanpa pendidikan yang baik maka suatu negara akan mengalami krisis sumber daya manusia sehingga perlunya pendidikan untuk menunjang keberhasilan dan berkembangnya suatu negara untuk mencapai tujuan yang ingin diraih. Pendidikan yang baik sesuai dengan urutan jenjang sekolah dari sekolah dasar (SD), Sekolah Menengah Pertama (SMP), Sekolah Menengah Atas (SMA). Ketiga jenjang ini wajib untuk di dapatkan oleh anak – anak yang berada pada suatu negara khususnya di Indonesia. Dari ketiga jenjang tersebut Sekolah Menengah Pertama (SMP) memiliki peranan yang sangat penting karena berada di tengah – tengah antara Sekolah Dasar (SD) dan Sekolah Menengah Atas (SMA) oleh karena itu untuk meningkatkan *kualitas* dan *kuantitas* Sekolah Menengah Pertama (SMP) perlunya untuk mengembangkan suatu aplikasi yang dapat di gunakan untuk memprediksi nilai UNAS

siswa Sekolah Menengah Pertama (SMP).

1.2 Rumusan Masalah

Penelitian ini merumuskan beberapa masalah, yakni:

1. Bagaimana hasil prediksi nilai UNAS siswa di suatu Sekolah Menengah Pertama?
2. Bagaimana mengimplementasikan metode *K – Means Clustering* dan *Regresi Linier Sederhana* dalam menghasilkan prediksi nilai UNAS siswa di suatu Sekolah Menengah Pertama ?
3. Bagaimana hasil prosentase tingkat akurasi prediksi nilai UNAS siswa dengan metode *RMSE (Root Means Squared Error)*?

1.3 Batasan Masalah

Agar penelitian ini tidak meluas dan dapat terarah sesuai dengan tujuan penelitian, maka permasalahan di batasi pada :

1. Aplikasi yang di bangun hanya untuk memprediksi nilai UNAS siswa di suatu Sekolah Menengah Pertama.

2. Data set yang di gunakan data nilai siswa SMP Negeri 2 Sumber Jambe Tahun 2012 kelas 3 semester 2 yaitu nilai ulangan harian, UTS dan UAS.
3. Proses untuk mengelompokkan data digunakan metode *K – Means Clustering* serta menghasilkan prediksi nilai UNAS menggunakan metode *Regresi Linier Sederhana* dan untuk mendapatkan nilai akurasi yang baik di gunakan metode *RMSE(Root Means Squared Error)*.

1.4 Tujuan Penelitian

Tujuan yang ingin di capai dalam penelitian ini yaitu :

1. Mengetahui hasil prediksi nilai UNAS siswa di suatu Sekolah Menengah Pertama.
2. Mengimplementasikan metode *K-Means Clustering* untuk mengelompokkan suatu data serta *Regresi Linier Sederhana* untuk menghasilkan prediksi nilai UNAS siswa di suatu Sekolah Menengah Pertama.

3. Mengetahui hasil tingkat *akurasi* prediksi UNASsiswa di suatu Sekolah Menengah Pertama dengan *RMSE(Root Means Squared Error)*.

1.5 Manfaat Penelitian

Beberapamanfaat yang diharapkan dari penelitian ini antara lain:

1. Menghasilkan suatu prediksi nilai UNASsiswa sebagai bahan evaluasi guru di Sekolah Menengah pertama.
2. Menghasilkan suatu perangkat lunak yang dapat memprediksi nilai UNASsiswa di suatu Sekolah Menengah Pertama.
3. Menghasilkan *prosentase* tingkat *akurasi* prediksi nilai UNASsiswa di suatu Sekolah Menengah Pertama.

TINJAUAN PUSTAKA

2.1 Pengertian K – Means

Clustering merupakan salah satu metode *eksplorasi* data yang digunakan dalam mencari pola yang ada pada suatu dataset. Pada umumnya pola tersebut dapat dilihat dari kesamaan sifat, *karakteristik*,

atau ciri dari *record-record* pada dataset. Salah satu metode *clustering* adalah *K - Means Clustering*.

K-means merupakan salah satu metode *clustering* non hirarki yang berusaha mempartisi data yang ada kedalam bentuk satu atau lebih *cluster*. Metode ini mempartisi data kedalam *cluster* sehingga data yang memiliki *karakteristik* yang sama dikelompokkan kedalam satu *cluster* yang sama dan data yang mempunyai *karakteristik* yang berbeda di kelompokkan kedalam *cluster* yang lain. Secara umum *algoritma* dasar dari *K-Means Clustering* adalah sebagai berikut :

1. Tentukan jumlah *cluster*
2. Alokasikan data kedalam *cluster* secara *random*
3. Hitung *centroid*/rata-rata dari data yang ada di masing-masing *cluster*
4. Alokasikan masing-masing data ke *centroid*/rata-rata terdekat
5. Kembalike Step 3, apabilamasihada data yang berpindah *cluster* atau apabila perubahan nilai *centroid*,ada yang di atas nilai *threshold* yang ditentukan atau apabila perubahan nilai pada *objective function* yang digunakan di

atas nilai *threshold* yang ditentukan. *Distance space* digunakan untuk menghitung jarak antara data dan centroid. Adapun persamaan yang dapat digunakan salah satunya yaitu *Euclidean Distance Space*. *Euclidean distance space* sering digunakan dalam perhitungan jarak, hal ini dikarenakan hasil yang diperoleh merupakan jarak terpendek antara dua titik yang diperhitungkan. Adapun persamaannya adalah sebagai berikut (Muliadinata, 2013) :

$$De = \sqrt{(Xi - Si)^2 + (Yi - ti)^2} \dots\dots\dots$$

Dimana :

De = jarak antar objek

Xi = Jarak objek Xi

Si = koordinat *cluster* Si

Yi = Jarak objek Yi

ti = koordinat *cluster* ti

Contoh :

Diberikan data nilai dari 5 siswa sebagai berikut kemudian jadikan data tersebut menjadi 2 cluster dengan menggunakan metode K – Means Clustering :

Tabel 2.1 Nilai Matematika dan Nilai Fisika Siswa

No	Nama	Nilai Mata Pelajaran	
		Matematika	Fisika
1	Anis Fadilah	79	75
2	Didik Mulyana	84	76
3	Eka Dwi Utami	77	84
4	Fandik Lasmana	78	86
5	Putri Khadijah	82	82

Penyelesaian :

1. Tentukan pusat awal cluster 'Centroid'

Untuk penentuan awal diasumsikan :

- Diambil data ke 3 sebagai pusat cluster ke 1 : (77 , 84)
- Diambil data ke 5 sebagai pusat cluster ke 2 : (82 , 82)

2. Perhitungan jarak pusat cluster

Untuk mengukur jarak antara data dengan pusat cluster digunakan Euclidian distance, kemudian akan di dapatkan matrik jarak sebagai berikut :

$$De = \sqrt{(Xi - Si)^2 + (Yi - ti)^2}$$

- Perhitungan Jarak dari data ke 1 terhadap pusat cluster :

$$C1 = \sqrt{(77 - 79)^2 + (84 - 75)^2}$$

$$= 11$$

$$C2 = \sqrt{(82 - 79)^2 + (82 - 75)^2}$$

$$= 10$$

- Perhitungan Jarak dari data ke 2 terhadap pusat cluster :

$$C1 = \sqrt{(77 - 84)^2 + (84 - 76)^2}$$

$$= 15$$

$$C2 = \sqrt{(82 - 84)^2 + (82 - 76)^2}$$

$$= 8$$

Perhitungan seterusnya sampai jarak dari data ke 5 terhadap pusat cluster, sehingga hasil perhitungan jarak selengkapnya pada iterasi 1 adalah :

Tabel 2.2 Nilai Hasil Cluster Iterasi 1

Data ke 1	Data ke 2	Data ke 3	Data ke 4	Data ke 5	Cluster
11	15	0	3	7	C1
10	8	7	8	0	C2

Baris pertama menunjukkan nilai jarak data terhadap titik pusat cluster pertama, baris kedua menunjukkan nilai jarak data terhadap pusat cluster kedua.

Pengelompokkan data didasarkan pada nilai terkecil jadi di dibandingkan antara nilai cluster 1 dan cluster 2 mana yang lebih kecil nilainya, disini di misalkan 0 berarti tidak masuk

cluster, jika 1 berarti masuk cluster 1 atau cluster 2 :

Tabel 2.3 Pengelompokan Cluster

Iterasi 1

Dat a ke 1	Dat a ke 2	Dat a ke 3	Dat a ke 4	Dat a ke 5	Cluste r
0	0	1	1	0	C1
1	1	0	0	1	C2

Penentuan pusat cluster baru diambil dari mana yang termasuk sebagai cluster 1 dan cluster 2 kemudian ditambahkan serta di bagi dengan banyaknya yang termasuk cluster 1 atau cluster 2 :

$$C1.1 = 77 + 78 / 2 = 77,5$$

$$C1.2 = 84 + 86 / 2 = 85$$

$$C2.1 = 79 + 84 + 82 / 3 = 81,6$$

$$C2.2 = 75 + 76 + 82 / 3 = 77,6$$

Pengulangan langkah ke 2 hingga posisi data tidak mengalami perubahan dengan menggunakan pusat cluster yang baru C1 (78 , 85) dan C2 (82 , 78) pada iterasi ke 2 yaitu di dapat hasil :

Tabel 2.4 Nilai Hasil Cluster Iterasi 2

Dat a ke 1	Dat a ke 2	Dat a ke 3	Dat a ke 4	Dat a ke 5	Cluste r
11	15	2	1	7	C1
6	4	11	11	4	C2

Jadi pengelompokan datanya pada iterasi 2 yaitu :

Tabel 2.5 Pengelompokan Cluster

Iterasi 2

Dat a ke 1	Dat a ke 2	Dat a ke 3	Dat a ke 4	Dat a ke 5	Cluste r
0	0	1	1	0	C1
1	1	0	0	1	C2

Karena pada iterasi 1 = iterasi 2 dimana anggota yang sama, maka tidak perlu dilakukan iterasi lagi dan sampai disini hasil clustering sudah mencapai stabil dan konvergen.

Hasil clustering adalah :

Cluster 1 : Eka dwi utami dan Fandik Lasmana

Cluster 2 : Anis Fadilah , Didik Mulyana , Putri Khadijah

2.2 SMP (Sekolah Menengah Pertama)

Sekolah menengah pertama disingkat SMP, dalam bahasa inggris yaitu *Junior High School* adalah jenjang pendidikan dasar pada pendidikan formal di Indonesia setelah lulus sekolah dasar atau sederajat (Wikipedia, 2014). Sekolah menengah pertama juga merupakan termasuk wajib belajar bagi setiap warga negara berusia 7 – 15 tahun di

Indonesia (Rizky, 2012). Sekolah menengah pertama ditempuh dalam waktu 3 tahun, mulai dari kelas 7 sampai kelas 9. Pada tahun ajaran 1994/1995 hingga 2003/2004, sekolah ini pernah disebut sebagai sekolah lanjutan tingkat pertama (SLTP). Murid kelas 9 diwajibkan untuk mengikuti Ujian Nasional dahulu disebut dengan EBTRANAS yang mempengaruhi kelulusan siswa. Lulusan sekolah menengah pertama dapat melanjutkan pendidikan ke sekolah menengah atas atau sekolah menengah kejuruan. Sekolah menengah pertama diselenggarakan oleh pemerintah maupun swasta. Sejak diberlakukannya otonomi daerah pada tahun 2001, pengelolaan sekolah menengah pertama negeri di Indonesia yang sebelumnya berada di bawah Departemen Pendidikan Nasional, kini menjadi tanggung jawab pemerintah daerah kabupaten atau kota.

(Wikipedia, 20014).

2.3 Regresi Linier

2.3.1 Pengertian Regresi Linier

Regresi Linier adalah suatu metode *statistik* yang dipergunakan

untuk mengetahui pengaruh antara satu atau beberapa variabel terhadap satu buah variabel. Variabel yang mempengaruhi sering disebut variabel bebas, variabel *independen* atau variabel penjelas. Variable yang dipengaruhi sering disebut dengan variabel terikat atau variabel *dependen*. *Regresi linier* hanya dapat digunakan pada skala *interval* dan *ratio*. Secara umum regresi linier terdiri dari dua yaitu *regresi linier* sederhana yaitu dengan satu buah variabel bebas dan satu buah variabel terikat. Serta *regresi linier* berganda dengan beberapa variable bebas dan satu buah variable terikat (Thalib, 2013). Analisis *regresi* adalah analisis yang mengukur pengaruh variabel bebas terhadap variabel terikat (Fadli, 2009).

2.3.2 Regresi Linier Sederhana

Regresi Linier Sederhana suatu model *statistik* yang digunakan untuk mengetahui pengaruh antara satu buah variabel bebas dengan satu buah variabel terikat. Di bawah ini rumus dari *regresi linier* sederhana :

$$Y = a + bx \dots\dots\dots$$

Persamaan diatas adalah rumus dari persamaan *regresi linier* sederhana. Y adalah variabel tak bebas, a adalah *koefisien intersep*, b adalah kemiringan dan x adalah variabel bebas. Rumus untuk b dan a adalah :

$$b = \frac{N\sum XY - (\sum X)(\sum Y)}{N\sum X^2 - (\sum X)^2}$$

$$a = \frac{\sum Y}{N} - b \frac{\sum X}{N}$$

Dalam *regresi linier* sederhana juga ada yang disebut dengan *koefisien* korelasi yang menunjukkan bahwa nilai suatu variabel bergantung pada perubahan nilai variabel yang lain. Rumus untuk menghitung *koefisien* korelasi adalah sebagai berikut (Makridakis, 1999) :

$$r = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{N\sum X^2 - (\sum X)^2} \sqrt{N\sum Y^2 - (\sum Y)^2}}$$

Contoh : Seorang Engineer ingin mempelajari Hubungan antara Suhu Ruangan dengan Jumlah Cacat yang diakibatkannya, sehingga dapat memprediksi atau meramalkan jumlah cacat produksi jika suhu

Tanggal	Rata – Rata Suhu Ruangan	Jumlah Cacat
1	24	10
2	22	5
3	21	6
4	20	3
5	22	6

ruangan tersebut tidak terkendali. Engineer tersebut kemudian mengambil data selama 5 hari terhadap rata-rata (mean) suhu ruangan dan Jumlah Cacat Produksi. Seperti tabel di bawah ini :

Langkah pertama : penentuan tujuan yaitu memprediksi jumlah cacat produksi jika suhu ruangan tidak terkendali. Langkah kedua : identifikasi variabel penyebab dan akibat yaitu variabel faktor penyebab X = Suhu ruangan, variabel akibat Y = jumlah cacat produksi. Langkah kedua yaitu menghitung X², Y², XY dan total dari masing –

masingnya. Berikut ini merupakan tabel yang telah dilakukan perhitungan dari X^2 , Y^2 , XY dan totalnya.

Tabel 2.6 Rata – Rata Suhu

Tanggal	Rata – Rata Suhu Ruang (X)	Jumlah Cacat (Y)	X^2	Y^2	XY
1	24	10	576	100	240
2	22	5	484	25	110
3	21	6	441	36	126
4	20	3	400	9	60
5	22	6	484	36	132
Total	109	30	2385	206	668

Langkah ketiga yaitu menghitung a dan b berdasarkan rumus regresi linier sederhana :

$$b = \frac{N \sum XY - (\sum X)(\sum Y)}{N \sum X^2 - (\sum X)^2}$$

$$b = \frac{(5 \cdot 668) - (109 \cdot 30)}{(5 \cdot 2385) - (109)^2}$$

$$= 1,59$$

Tabel 2.7 Proses Regresi Linier

$$a = \frac{\sum Y}{N} - b \frac{\sum X}{N}$$

$$a = \frac{30}{5} - (1,6) \frac{109}{5}$$

$$= 6 - 34,88$$

$$= -28,88$$

$$= -29$$

Langkah keempat yaitu model persamaan regresi :

$$Y = a + bx$$

$$Y = -29 + 1,6X$$

Langkah terakhir :

Melakukan prediksi atau peramalan terhadap variabel faktor penyebab atau variabel akibat yaitu memprediksi jumlah cacat produksi jika suhu dalam keadaan tinggi (variabel X) sebesar 30°C yaitu :

$$Y = -29 + 1,6X$$

$$= -29 + (1,6 \cdot 30)$$

$$= 19$$

Jadi jika suhu ruangan mencapai 30°C maka akan di prediksi terdapat 19 unit cacat yang dihasilkan oleh produksi.

2.4 RMSE (Root Mean Squared Error)

Metode rata –rata bergerak ini dilakukan untuk lebih memperbaiki hasil dari metode trend yang dinilai cukup kasar dan cukup beresiko penyimpangan dari hasil trend. Metode rata – rata bergerak sesuai dengan namanya bergerak dilakukan pengelompokan periode waktu dihitung rata – ratanya menurut pengelompokan waktu yang dinilai dapat mewakili sifat data yang akan dihitung. Misalnya apakah dihitung per tiga, empat atau lima bulanan dalam suatu satuan waktu yang akan kita pilih. Ini tergantung dari sifat data maupun kepentingan peramalan terhadap data yang tersedia. Sifat data yang akan dihitung rata – ratanya setidaknya dilatarbelakangi oleh kondisi umum yang normal, dan tidak terjadi gejolak yang tak teratur yang mencerminkan ketidaknormalan kondisi sehingga data muncul terlalu *berfluktuasi* yang sangat tinggi.

Kepentingan peramalan yang akan dipakai apakah dalam waktu jangka pendek nilai ramalan yang diketahui atau dalam jangka atau rentan waktu yang lebih panjang.

Jika menggunakan rentan waktu yang lebih pendek maka hasil rata – rata bergerak yang kita peroleh akan lebih mendekati kondisi sifat data yang sebenarnya dan rata – rata yang ditemukan terdistribusi atau tersebar pada kelompok data faktual. Sedangkan jika satuan waktu yang lebih panjang, rata – rata yang akan kita peroleh akan lebih mewakili sejumlah data yang lebih banyak dan beraneka ragam *fluktuasinya*, sehingga rata – rata bergerak lebih tersebar dan kurang mewakili fakta sifat data yang tersebar tersebut. Terkecuali sifat data lebih *homogen* dan tidak terlalu *fluktuatif*. Angka *deviasi* dapat diperoleh dengan cara mengurangi angka *observasi* dengan angka rata – rata bergerak. Sedangkan *deviasi* kuadrat adalah mengkuadratkan *deviasi*. Jumlah *deviasi* kuadrat ini akan dipergunakan untuk menghitung penyimpangan atau tingkat *error* istilah disini adalah *root mean square error* (RMSE) setiap

pengelompokkan rata – rata bergerak ini (Cici, 2012). Rumus untuk RMSE seperti berikut ini :

$$RMSE = \frac{\sqrt{\sum (A_t - F_t)^2}}{n}$$

Keterangan :

RMSE = *Root Mean Square*

Error

A_t = Angka observasi

F_t = Rata – rata bergerak

yang dihasilkan perkelompok waktu

Contoh :

Tentukan RMSE dari ramalan penjualan pada januari 2007

Tabel 2.8 Ramalan Penjualan Tahun 2006

Bulan2006	Penjualan	Peramalan
1	20	21
2	22	20,5
3	23	21,25
4	24	22,13
5	18	23,07
6	23	20,54
7	19	21,77
8	17	20,39
9	22	18,70
10	23	20,35
11	18	21,68
12	23	19,84

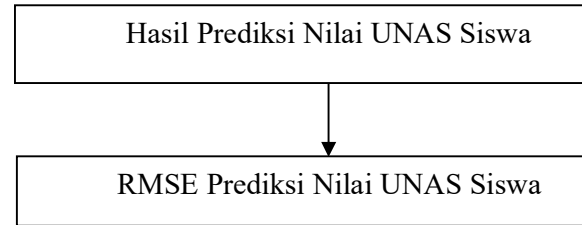
Penyelesaian :

Tabel 2.9 Proses Perhitungan RMSE

Bulan	Penjualan	Peramalan	A – F	(A – F) ²
2006			7	
1	20	21	-1	1
2	22	20,5	1,5	2,25
3	23	21,25	1,7	3,062
4	24	22,13	1,8	3,496
5	18	23,07	-	25,70
6	23	20,54	2,4	6,051
7	19	21,77	-	7,672
8	17	20,39	-	11,49
9	22	18,70	3,3	10,89
10	23	20,35	2,6	7,022
11	18	21,68	-	13,54
12	23	19,84	3,1	9,985

Jumlah				102,1223
--------	--	--	--	----------

$$\text{Jadi RMSE} = \sqrt{\frac{102,1223}{12}} = 2,92$$



BAB III METODE PENELITIAN

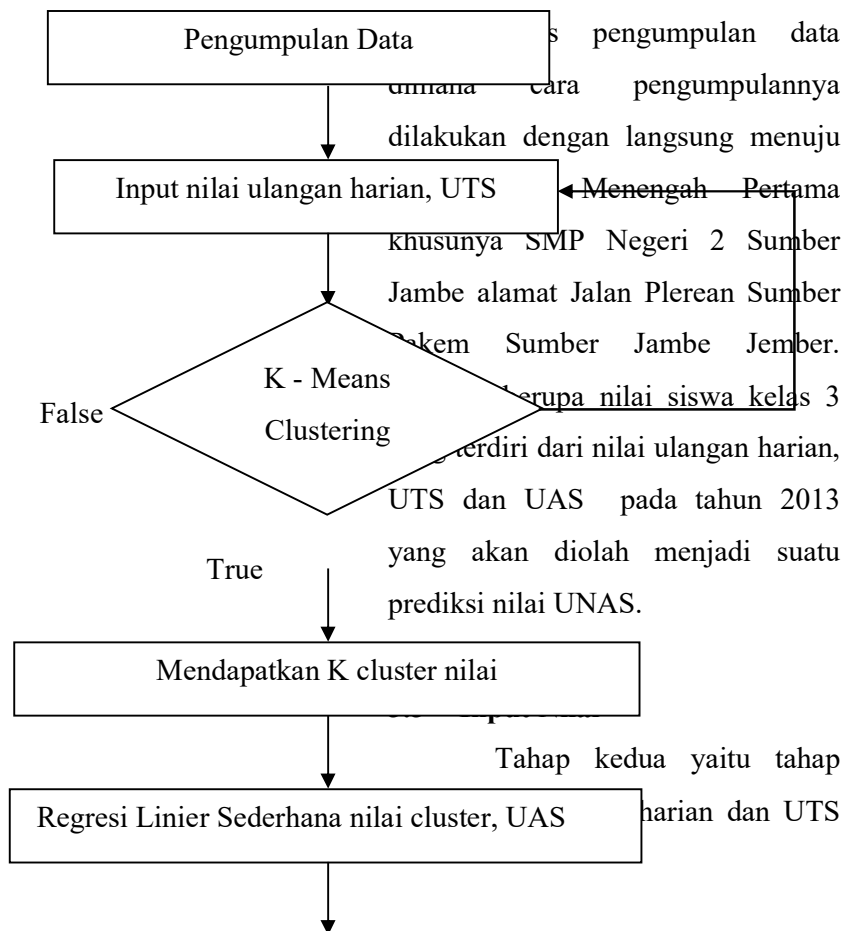
Gambar 3.1 Blok Diagram Rancangan Penelitian Prediksi Nilai UNAS Siswa

3.1 Rancangan Penelitian

Rancangan penelitian dalam tugas akhir ini mengikuti blok diagram berikut.

3.2 Pengumpulan Data

Tahap awal di dalam penelitian prediksi kelulusan siswa



dari 4 mata pelajaran yaitu matematika, bahasa indonesia, bahasa inggris dan IPA. Kedua

Cluster 2 (88,79) , Cluster 3 (87,89) yang di proses di Microsoft Excell sehingga di dapat hasil yaitu :

parameter tersebut di proses dengan menggunakan metode K - Means berdasarkan K cluster sebuah kondisi apabila yang di inputkan ulangan harian dan UTS tersebut akan di berdasarkan K cluster akan berbentuk K Apabila data nilai maka proses kembali nilai ulangan harian dan UTS Contoh pengelompokan berdasarkan K - Means dengan perumusannya yaitu

PROSES K - MEANS CLUSTERING MATA PELAJARAN BAHASA INDONESIA													
ITERASI 1 (PUSAT C1(91,85), PUSAT C2 (90,80), PUSAT C3(92,85))													
NO	NAMA	NILAI ULANGAN HARIAN	UTS	PUSAT (C1 X)	PUSAT (C1 Y)	PUSAT (C2 X)	PUSAT (C2 Y)	PUSAT (C3 X)	PUSAT (C3 Y)	NILAI CLUS TER 1	NILAI CLUS TER 2	NILAI CLUS TER 3	
1	Agustina Fadilatul	93.83	85	91	85	90	80	92	85	2.83	6.30	1.83	
2	Asanah	91.08	85	91	85	90	80	92	85	0.08	5.12	0.92	
3	Apul Pankoini	92.92	80	91	85	90	80	92	85	5.36	2.92	5.08	
4	Fiqih Firmansyah	92.5	80	91	85	90	80	92	85	5.22	2.50	5.02	
5	Hafiyatul Hasanah	90.83	85	91	85	90	80	92	85	0.17	5.07	1.17	
6	Ilham Hamid	88.58	85	91	85	90	80	92	85	2.42	5.20	3.42	
7	Imam Syaiful	90.75	85	91	85	90	80	92	85	0.25	5.06	1.25	
8	Imron Rosidi	90.5	85	91	85	90	80	92	85	0.50	5.02	1.50	
9	Imrotul Hasanah	92.25	85	91	85	90	80	92	85	1.25	5.48	0.25	
10	Indah Purnama Sari	90.25	85	91	85	90	80	92	85	0.75	5.01	1.75	
11	Matlubatul Hasanah	91.92	85	91	85	90	80	92	85	0.92	5.36	0.08	
12	Misbahul Arifin	88.83	78	91	85	90	80	92	85	7.33	2.32	7.68	
13	Moch. Wakil	89.42	80	91	85	90	80	92	85	5.24	0.58	5.63	
14	Mochamad Yahya F.	94.5	85	91	85	90	80	92	85	3.50	6.73	2.50	
15	Mohamad Sudarsoino	91	85	91	85	90	80	92	85	0.00	5.10	1.00	
16	Mohammad Rosi Amir	91.17	80	91	85	90	80	92	85	5.00	1.17	5.07	
17	Muhammad Lutfi	88.17	80	91	85	90	80	92	85	5.75	1.83	6.30	
18	Muhammad Subli	91.08	80	91	85	90	80	92	85	5.00	1.08	5.08	
19	Riskiatal Hasanah	87.5	80	91	85	90	80	92	85	6.10	2.50	6.73	
20	Rofiatul Ardawiyah	90	80	91	85	90	80	92	85	5.10	0.00	5.39	
21	Ruly Rifky Hadi	90.08	85	91	85	90	80	92	85	0.92	5.00	1.92	
22	Salim Umar	89.33	85	91	85	90	80	92	85	1.67	5.04	2.67	
23	Sayid Abdullah	90.25	90	91	85	90	80	92	85	5.06	10.00	5.30	
24	Siti Nur'aini	89.75	85	91	85	90	80	92	85	1.25	5.01	2.25	
25	Siti Rofikah	89.67	85	91	85	90	80	92	85	1.33	5.01	2.33	
26	Sri Wahyuni	90.5	85	91	85	90	80	92	85	0.50	5.02	1.50	
27	Sulistiyowati	92	85	91	85	90	80	92	85	1.00	5.39	0.00	
28	Sumiyati	91.75	80	91	85	90	80	92	85	5.06	1.75	5.01	
29	Sunaryo	89.42	80	91	85	90	80	92	85	5.24	0.58	5.63	
30	Suzul Saeh	89.75	85	91	85	90	80	92	85	1.25	5.01	2.25	

De

$$\sqrt{(X_i - S_i)^2 + (Y_i - t_i)^2}$$

- Dimana :
- De = jarak antara objek Xi dan cluster Si
 - Xi = Jarak objek Xi
 - Si = koordinat cluster Si
 - Yi = Jarak objek Yi
 - ti = koordinat cluster ti

Pada Iterasi 1 dengan Cluster Di bawah ini merupakan proses K - Means dengan Cluster 1 (91,85) , Cluster 2 (90,80) dan Cluster 3 (87,89)

Cluster 3(92,85) di gunakan rumus yaitu :

$$\text{Cluster 1} = \sqrt{(91 - 93,83)^2 + (85 - 85)^2} = 2,83$$

$$\text{Cluster 2} = \sqrt{(90 - 93,83)^2 + (80 - 85)^2} = 6,3$$

$$\text{Cluster 3} = \sqrt{(92 - 93,83)^2 + (85 - 85)^2} = 1,83$$

Perhitungan tersebut dilakukan sampai mencapai data ke 30 seperti yang ada pada tabel diatas, di dalam pengelompokan cluster dibandingkan nilai yang terkecil dari masing – masing sehingga yang termasuk dalam cluster 1 yaitu yang berwarna kuning , kemudian Cluster 2 yang berwarna biru dan Cluster 3 yaitu yang berwarna coklat. Iterasi akan berhenti apabila data hasil perhitungan cluster yang baru dengan yang lama sama atau tetap maka cluster di hentikan. Apabila cluster baru dan lama masih berbeda maka iterasi terus di lakukan.

3.4 Perhitungan Regresi Linier Sederhana

Pada masing – masing clusternya kemudian dimodelkan dengan persamaan Regresi Linier Sederhana dengan rumus :

$$Y = a + bx$$

Keterangan:

Y = Variabel tak bebas

a = Koefisien

b = Kemiringan

x = Variabel bebas

$$b = \frac{N \sum XY - (\sum [X]) (\sum [Y])}{N \sum X^2 - (\sum [X])^2}$$

$$a = \frac{\sum Y}{N} - b \frac{\sum X}{N}$$

PROSES REGRESI LINIER SEDERHANA MATA PELAJARAN BAHASA INDONESIA							
NO	NAMA	NILAI	UAS	(Y + X) / 2	Y . X	X ²	Y Prediksi
		CLUSTER	(X)	RMSE			
		(Y)					
1	Agustina Fadilatul f.	93.85	75	84.425	7038.75	5625	91.69
2	Asanah	90.85	70	80.425	6359.5	4900	91.26
3	Ayyul Farikoinih	90.8	72	81.4	6537.6	5184	91.43
4	Fiqih Firmansyah	90.8	65	77.9	5902	4225	90.83
5	Hafiyatul Hasanah	90.85	80	85.425	7268	6400	92.11
6	Ilham Rahmat	90.85	65	77.925	5905.25	4225	90.83
7	Imam Syafi'i	90.85	70	80.425	6359.5	4900	91.26
8	Imron Rosidi	90.85	65	77.925	5905.25	4225	90.83
9	Imrotul Hasanah	93.85	70	81.925	6569.5	4900	91.26
10	Indah Purnama Sari	90.85	75	82.925	6813.75	5625	91.69
11	Matlubatul Hasanah	93.85	70	81.925	6569.5	4900	91.26
12	Misbahul Arifin	90.8	66	78.4	5992.8	4356	90.92
13	Moch. Wakil	90.8	65	77.9	5902	4225	90.83
14	Mochamad Yahya F.	93.85	80	86.925	7508	6400	92.11
15	Mohamad Sudarsoino	90.85	80	85.425	7268	6400	92.11
16	Mohammad Rosi Amin	90.8	66	78.4	5992.8	4356	90.92
17	Muhammad Lutfi	90.8	65	77.9	5902	4225	90.83
18	Muhammad Subli	90.8	70	80.4	6356	4900	91.26
19	Riskiatul Hasanah	90.8	70	80.4	6356	4900	91.26
20	Rofiatul Ardawiyah	90.8	70	80.4	6356	4900	91.26
21	Ruly Rifky Hadi	90.85	75	82.925	6813.75	5625	91.69
22	Salim Umar	90.85	70	80.425	6359.5	4900	91.26
23	Sayid Abdullah	90.85	72	81.425	6541.2	5184	91.43
24	Siti Nur'aini	90.85	70	80.425	6359.5	4900	91.26
25	Siti Rofikah	90.85	75	82.925	6813.75	5625	91.69
26	Sri Wahyuni	90.85	70	80.425	6359.5	4900	91.26
27	Sulistiyowati	93.85	75	84.425	7038.75	5625	91.69
28	Sumiyati	90.8	70	80.4	6356	4900	91.26
29	Sunaryo	90.8	70	80.4	6356	4900	91.26
30	Tanwirotul Af'ideh	90.85	70	80.425	6359.5	4900	91.26
	SUM	2739.95	2126.00	2432.98	194219.65	151230.00	2740.01

b

$$\frac{N \sum XY - (\sum [X]) (\sum [Y])}{N \sum X^2 - (\sum [X])^2}$$

b

$$= \frac{(30 \cdot 194219,65) - (2126 \cdot 2740,01)}{(30 \cdot 151230) - (2126)^2}$$

$$= \frac{1455,8}{17024} = 0,086$$

$$a = \frac{\sum Y}{N} - b \frac{\sum X}{N}$$

$$a = \frac{2740,01}{30} -$$

$$\left(0,086 \cdot \left(\frac{2126}{30} \right) \right)$$

$$= 91,33 - 6,095$$

$$= 85,24$$

Dari perhitungan di atas di dapat persamaan untuk menghasilkan Y prediksi yaitu

$$Y = a + bx$$

$$Y = 85,24 + (0,086 * 75)$$

$$Y = 91,69$$

Persamaan di atas akan di proses yaitu mengganti variabel X dengan nilai UAS setiap siswa sehingga di dapatkan suatu prediksi nilai siswa yang kemudian akan di lakukan

perhitungan untuk mengetahui nilai akurasi prediksi tersebut

3.5 RMSE Prediksi Kelulusan

Pada tahap terakhir ini dilakukan proses RMSE (Root Mean Squared Error). Hasil dari prediksi tingkat kelulusan siswa di bandingkan dengan nilai parameter inputan siswa, kemudian di hitung berapa tingkat akurasinya. Semakin kecil nilai RMSE nya maka semakin akurat prediksinya, begitu juga sebaliknya. Menggunakan rumus :

$$RMSE = \sqrt{\frac{\sum (A_t - F_t)^2}{n}}$$

Dimana :

RMSE = Root Mean Square Error

A_t = Angka observasi

F_t = Rata-rata bergerak yang dihasilkan per kelompok waktu

n = Banyaknya data yang diobservasi

$$RMSE = \sqrt{\frac{(2432,98 - 2740,01)^2}{30}}$$

$$RMSE = 56,0457$$

Jadi untuk menghitung tingkat akurasinya yaitu akar dari jumlah data yang di prediksi di kurangi jumlah prediksi yang di hasilkan kemudian kuaratkan di bagi dengan banyaknya jumlah data dihasilkan tingkat akurasi error untuk mata pelajaran Bahasa Indonesia yaitu 8,61.

HASIL DAN PEMBAHASAN

4.1 Implementasi Proses Peramalan

Dalam bab ini akan dipaparkan proses prediksi nilai UNAS siswa SMP dengan metode *K-means clustering* dan regresi linier sederhana, dan menghitung kesalahan prediksi dengan metode *Root Mean Squared Error* (RMSE).

4.2 Proses Perhitungan *K-means*.

Berikut ini akan dijabarkan beberapa proses perhitungan *k-means*, regresi linier sederhana dan kesalahan prediksi dengan *Root Mean squared Error* (RMSE).

4.2.1 Prediksi *K-means*

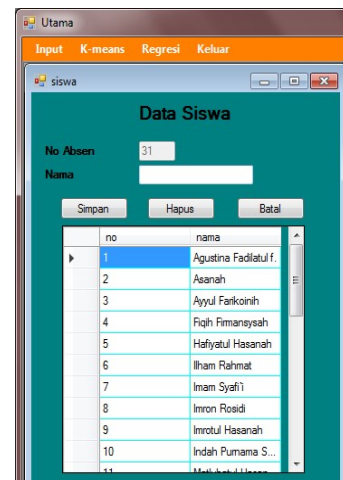
Data yang digunakan adalah data nilai ulangan harian, dan UTS pada mata pelajaran Bahasa Indonesia dan menggunakan 3 pusat. Centroid 1 mengambil nilai ulangan dan UTS data ke 15, centroid 2 mengambil data ke 20, dan centroid 3 mengambil data ke 27.

13	Moch. Wakil	89.42
14	Mochamad Yahya F.	94.5
15	Mohamad Sudarsoino	91
16	Mohammad Rosi Amin	91.17
17	Muhammad Lutfi	88.17
18	Muhammad Subli	91.08
19	Riskiatal Hasanah	87.5
20	Rofiatul Ardawiyah	90
21	Ruly Rifky Hadi	90.08
22	Salim Umar	89.33
23	Sayid Abdulllah	90.25
24	Siti Nur'aini	89.75
25	Siti Rofikah	89.67
26	Sri Wahyuni	90.5
27	Sulistiyowati	92
28	Sumiyati	91.75
29	Sunaryo	89.42
30	Tanwirotul Af'ideh	89.75

Tabel 4.1 Nilai Bahasa Indonesia

No	Nama	UH	UTS
1	Agustina Fadilatul f.	93.83	85
2	Asanah	91.08	85
3	Ayyul Farikoinih	92.92	80
4	Fiqih Firmansyah	92.5	80
5	Hafiyatul Hasanah	90.83	85
6	Ilham Rahmat	88.58	85
7	Imam Syafi'i	90.75	85
8	Imron Rosidi	90.5	85
9	Imrotul Hasanah	92.25	85
10	Indah Purnama Sari	90.25	85
11	Matlubatul Hasanah	91.92	85
12	Misbahul Arifin	88.83	78

Gambar 4.1 Tampilan Data Input Siswa



Gambar 4.2 Tampilan Input Nilai Siswa



Perhitungan *K-Means*

Iterasi 1 :

Centroid 1 (91,85) ;
centroid 2 (90,80), dan
centroid 3 (92,85)

$$De = \sqrt{(X_i - S_i)^2 + (Y_i - t_i)^2}$$

Nilai kluster 1 :

$$C1 = \sqrt{(91 - 93,83)^2 + (85 - 85)^2}$$

$$C1 = \sqrt{(-2,83)^2 + (0)^2}$$

$$C1 = \sqrt{8,0089}$$

$$C1 = 2,83$$

Nilai kluster 2 :

$$C2 = \sqrt{(90 - 93,83)^2 + (80 - 85)^2}$$

$$C2 = \sqrt{(-3,83)^2 + (-5)^2}$$

$$C2 = \sqrt{14,6689 + 25}$$

$$C2 = \sqrt{39,6689}$$

$$C2 = 6,3$$

$$C2 = 6,3$$

Nilai kluster 3 :

$$C3 = \sqrt{(92 - 93,83)^2 + (85 - 85)^2}$$

$$C3 = \sqrt{(-1,83)^2 + (0)^2}$$

$$C3 = \sqrt{3,3489}$$

$$C3 = 1,83$$

Perhitungan kluster dilakukan sampai seluruh data pada mata pelajaran Bahasa Indonesia dihitung.

Tabel 4.2 Iterasi 1

Bahasa

Indonesia

No	Nama	Kluster1	Kluster2	Kluster3	Posisi
1	Agustina Fadilatul f.	2.83	6.3	1.83	3
2	Asanah	0.08	5.12	0.92	1
3	Ayyul Farkoimih	5.36	2.92	5.08	2
4	Fiqih Firmansyah	5.22	2.5	5.02	2
5	Hafiyatul Hasanah	0.17	5.07	1.17	1
6	Ilham Rahmat	2.42	5.2	3.42	1
7	Imam Syaff'i	0.25	5.06	1.25	1
8	Imron Rosidi	0.5	5.02	1.5	1
9	Imrotul Hasanah	1.25	5.48	0.25	3
10	Indah Purnama Sari	0.75	5.01	1.75	1
11	Matubatul Hasanah	0.92	5.36	0.08	3
12	Misbahul Arifin	7.33	2.32	7.68	2
13	Moch. Wakil	5.24	0.58	5.63	2
14	Mochamad Yahya F.	3.5	6.73	2.5	3
15	Mohamad Sudarsoino	0	5.1	1	3
16	Mohammad Rosi Amin	5	1.17	5.07	2
17	Muhammad Lutfi	5.75	1.83	6.3	2
18	Muhammad Subli	5	1.08	5.08	2
19	Riskiatal Hasanah	6.1	2.5	6.73	2
20	Rofiatul Ardawiyah	5.1	0	5.39	3

21	Ruby Rizky Hadi	0.92	5	1.92	1
22	Salm Umar	1.67	5.04	2.67	1
23	Sayid Abdullah	5.06	10	5.3	1
24	Siti Nur'aini	1.25	5.01	2.25	1
25	Siti Roikah	1.33	5.01	2.33	1
26	Sri Wahyuni	0.5	5.02	1.5	1
27	Substiyowati	1	5.39	0	3
28	Sumiyati	5.06	1.75	5.01	2
29	Sunaryo	5.24	0.58	5.63	2
30	Tamwirotul Afideh	1.25	5.01	2.25	1

Iterasi 2 :

Menentukan

pusat baru di ambil dari data mana yang termasuk sebagai kluster 1, kluster 2 dan kluste 3 kemudian ditambahkan serta dibagi dengan banyaknya data

yang termasuk dalam kluster 1, kluster 2 atau kluster 3.

Pusat 1

$$C1.1 = (91,08 + 90,83 + 88,58 + 90,75 + 90,5 + 90,25 + 91 + 90,08 + 89,33 + 90,25 + 89,75 + 89,67 + 90,5 + 89,75) / 14$$

$$C1.1 = 1262,32 / 14$$

$$C1.1 = 90,17$$

$$C1.2 = (85 + 85 + 85 + 85 + 85 + 85 + 85 + 85 + 85 + 85 + 85 + 85 + 85 + 85 + 85) / 14$$

$$C1.2 = 1195 / 14$$

$$C1.2 = 85,36$$

Sehinggapusat 1 yang baru adalah 90,17 dan 85,36

Pusat 2

$$C2.1 = (92,92 + 92,5 + 88,83 + 89,42 + 91,17$$

$$\begin{aligned}
 &+ 88,17 + \\
 &91,08 + 87,5 \\
 &+ 90 + 91,75 \\
 &+ 89,42) / 11
 \end{aligned}$$

$$C2.1 = 992,76 / 11$$

$$C2.1 = 90,25$$

$$\begin{aligned}
 C2.2 &= (80 + 80 + 78 \\
 &+ 80 + 80 + 80 + 80 + 80 \\
 &+ 80 + 80 + 80) / 11
 \end{aligned}$$

$$C2.2 = 878 / 11$$

$$C2.2 = 79,82$$

Sehingga pusat 2 yang baru adalah 90,25 dan 79,82

Pusat 3

$$\begin{aligned}
 C3.1 &= (93,83 + 92,25 \\
 &+ 91,92 + 94,5 + 92) / 5
 \end{aligned}$$

$$C3.1 = 464,5 / 5$$

$$C3.1 = 92,9$$

$$\begin{aligned}
 C3.2 &= (85 + 85 + 85 \\
 &+ 85 + 85) / 5
 \end{aligned}$$

$$C3.2 = 425 / 5$$

$$C3.2 = 85$$

Sehingga pusat 2 yang baru adalah 92,5 dan 85

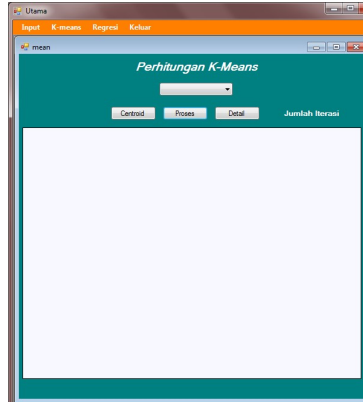
Setelah pusat baru terbentuk dilakukan proses perhitungan kluster lagi seperti pada iterasi 1 tetapi menggunakan pusat yang

baru. Proses iterasi dilakukan berulang-ulang hingga posisi kluster tidak mengalami perubahan.

Tabel 4.3 Iterasi 2 Bahasa Indonesia

No	Nama	Kluster1	Kluster2	Kluster3	Posisi
1	Agustina Fadhatul f	3,75	6,30	1,77	3,00
2	Asarah	1,05	5,26	1,34	1,00
3	Ayyul Farkoinh	6,08	2,65	4,34	2,00
4	Fiqih Fimansyah	5,89	2,23	4,30	2,00
5	Haftiyatul Hasanah	0,82	5,23	1,56	1,00
6	Ilham Rahmat	1,57	5,47	3,70	1,00
7	Imam Syafi'i	0,75	5,22	1,63	1,00
8	Imron Rosidi	0,55	5,20	1,86	1,00
9	Imrotul Hasanah	2,18	5,56	0,72	3,00
10	Indah Purnama Sari	0,41	5,20	2,09	1,00
11	Mathabatul Hasanah	1,86	5,45	0,77	3,00
12	Mishahul Arifin	7,49	2,31	7,14	2,00
13	Moch. Wakil	5,43	0,88	5,12	2,00
14	Mochamad Yahya F.	4,42	6,70	2,39	3,00
15	Mohamad Sudarsono	0,98	5,25	1,41	1,00
16	Mohammad Rosi Amin	5,49	0,92	4,41	2,00
17	Muhammad Luthi	5,72	2,12	5,89	2,00
18	Muhammad Subfi	5,47	0,83	4,43	2,00
19	Riskiatal Hasanah	5,98	2,78	6,37	2,00
20	Rofiatul Ardawiyah	5,39	0,34	4,82	2,00
21	Ruby Rifky Hadi	0,39	5,20	2,25	1,00
22	Salim Umar	0,86	5,29	2,97	1,00
23	Sayid Abdullah	4,62	10,20	6,04	1,00
24	Sri Nur'ani	0,52	5,23	2,57	1,00
25	Sri Roftikah	0,58	5,24	2,64	1,00
26	Sri Wahyuni	0,55	5,20	1,86	1,00
27	Sulisdiyowati	1,94	5,48	0,75	3,00
28	Sumiyati	5,63	1,49	4,31	2,00
29	Sunaryo	5,43	0,88	5,12	2,00
30	Tanwirotul Af'ideh	0,52	5,23	2,57	1,00

Gambar 4.3 Perhitungan K-Means



4.2.2 Regresi Linier sederhana

RMSE

Pada perhitungan linier sederhana data yang digunakan adalah data nilai kluster, dan UAS pada mata pelajaran Bahasa Indonesia, dengan Y adalah nilai kluster dan X adalah nilai UAS.

Perhitungan regresi linier sederhana :

$$Y = a + bx$$

$$b = \frac{N \sum XY - (\sum X)(\sum Y)}{N \sum X^2 - (\sum X)^2}$$

$$b = \frac{(30 * 194219,7) - (2739,95 * 2126)}{(30 * 151230) - ([2126])^2}$$

$$b = \frac{(5826589,5) - (5825133,7)}{(4536900) - (4519876)}$$

$$b = \frac{(5826589,5) - (5825133,7)}{(4536900) - (4519876)}$$

$$b = \frac{1455,8}{17024}$$

$$b = 0,086$$

$$a = \frac{\sum Y}{N} - b \frac{\sum X}{N}$$

$$a = \frac{2739,95}{30} -$$

$$0,086 \frac{2126}{30}$$

$$a = 85,27$$

Dari perhitungan di atas didapat persamaan untuk menghasilkan Y prediksi yaitu :

$$Y = a + bx$$

$$Y = 85,27 + (0,086 * 75)$$

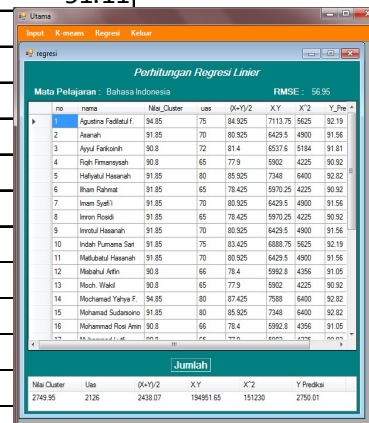
$$Y = 91,69$$

Proses perhitungan regresi linier dilakukan sampai semua data di proses.

Tabel 4.4 Regresi

Linier Sederhana

No	Nama	Nilai Cluster	Uas	Y Prediksi	RMSE :
1	Agustina Fadilatul f.	92.84	75	91.4	
2	Asanah	90.85	70	91.11	
3	Ayyul Farikoinih	90.8	72	91.23	
4	Fiqih Firmansyah	90.8	65	90.83	
5	Hafiyatul Hasanah	90.85	80	91.69	
6	Ilham Rahmat	90.85	65	90.83	
7	Imam Syafi'i	90.85	70	91.11	
8	Imron Rosidi	90.85	65	90.83	
9	Imrotul Hasanah	92.84	70	91.11	
10	Indah Purnama Sari	90.85	75	91.455,57	
11	Matlubatul Hasanah	92.84	70	91.11	
12	Misbahul Arifin	90.8	66	90.88	
13	Moch. Wakil	90.8	65	90.83	
14	Mochamad Yahya F.	92.84	80	91.69	
15	Mohamad Sudarsoino	90.85	80	91.69	
16	Mohammad Rosi Amin	90.8	66	90.88	
17	Muhammad Lutfi	90.8	65	90.83	
18	Muhammad Subli	90.8	70	91.11	
19	Riskiatul Hasanah	90.8	70		
20	Rofiatul Ardawiyah	90.8	70		
21	Ruly Rifky Hadi	90.85	75		
22	Salim Umar	90.85	70		
23	Sayid Abdullhah	90.85	72		
24	Siti Nur'aini	90.85	70		
25	Siti Rofikah	90.85	75		
26	Sri Wahyuni	90.85	70		
27	Sulistiyowati	92.84	75		
28	Sumiyati	90.8	70		
29	Sunaryo	90.8	70		
30	Tanwirotul Af'ideh	90.85	70		



4.3 Hasil K-means 4 Centroid dan Regresi Linier

**Sederhana Bahasa
Indonesia**

Centroid I : 5 Centroid II
: 12 Centroid III : 20
Centroid IV : 25

Tabel 4.5 K-means Iterasi4

No	Nama	Kluster1	Kluster2	Kluster3	Kluster4	Posisi
1	Agustina Fadilatul f.	2.115122917	8.011520178	5.942536715	3.998734796	1
2	Asanah	0.843544625	6.371704833	5.08819262	1.258147658	1
3	Ayyul Farikoinih	5.602955998	4.799856479	2.179819221	5.718062542	3
4	Fiqih Firmansyah	5.532117933	4.384315479	1.760143216	5.5028816	3
5	Hafiyatul Hasanah	1.058088562	6.261346146	5.077693706	1.011540066	4
6	Ilham Rahmat	3.227445777	5.681721179	5.517915952	1.269138805	4
7	Imam Syaifi	1.129825758	6.227738131	5.076930128	0.933007071	4
8	Imron Rosidi	1.3599596	6.12825877	5.082665517	0.689768057	4
9	Imrotul Hasanah	0.680883997	6.984606089	5.296282106	2.421535779	1
10	Indah Purnama Sari	1.59608082	6.037498562	5.100662505	0.453935844	4
11	Matlubatul Hasanah	0.506302012	6.796956885	5.211901719	2.092654342	1
12	Misbahul Arifin	8.040995953	1.489224257	2.711494817	6.888602393	2
13	Moch. Wakil	5.966564348	1.41960871	1.323775308	4.832714364	3
14	Mochamad Yahya F.	2.771596684	8.498365856	6.316738167	4.668119057	1
15	Mohamad Sudarsoino	0.910232043	6.335525936	5.083497841	1.179077782	1
16	Mohammad Rosi Amin	5.516702187	3.076435528	0.435311899	4.996317011	3
17	Muhammad Lutfi	6.560473266	0.666675	2.572688714	5.094753871	2
18	Muhammad Subli	5.527230519	2.988637742	0.347092744	4.973033566	3
19	Riskiatal Hasanah	6.950827738	0.942809042	3.242451042	5.351379397	2
20	Rofiatul Ardawiyah	5.762739586	1.950783318	0.745517572	4.817622201	3
21	Ruly Rifky Hadi	1.758835699	5.980966106	5.119841898	0.306670245	4
22	Salim Umar	2.488464312	5.784847064	5.269496063	0.538405027	4
23	Sayid Abdullah	4.765371371	10.86821308	10.08890424	5.201721792	1
24	Siti Nur'aini	2.078086073	5.883711036	5.172842202	0.203993397	4
25	Siti Rofikah	2.155967664	5.862688992	5.188770818	0.248399265	4
26	Siti Wahyuni	1.3599596	6.12825877	5.082665517	0.689768057	4
27	Sulistiyowati	0.535338107	6.841458584	5.230571037	2.172352022	1
28	Sumiyati	5.483911053	3.644821288	1.011391039	5.18146111	3
29	Sunaryo	5.966564348	1.41960871	1.323775308	4.832714364	3
30	Tanwirotul Af'ideh	2.078086073	5.883711036	5.172842202	0.203993397	4

**Tabel 4.6 Regresi Linier
Sederhana**

No	Nama	Nilai Cluster	Uas	Y Predikasi
1	Agustina Fadilatul f.	92.85	75	91.84
2	Asanah	92.85	70	91.35
3	Ayyul Farikoinih	91.8	72	91.54
4	Fiqih Firmansyah	91.8	65	90.85
5	Hafiyatul Hasanah	90.85	80	92.33
6	Ilham Rahmat	90.85	65	90.85
7	Imam Syaifi	90.85	70	91.35
8	Imron Rosidi	90.85	65	90.85
9	Imrotul Hasanah	92.85	70	91.35
10	Indah Purnama Sari	90.85	75	91.84
11	Matlubatul Hasanah	92.85	70	91.35
12	Misbahul Arifin	88.79	66	90.95
13	Moch. Wakil	91.8	65	90.85
14	Mochamad Yahya F.	92.85	80	92.33
15	Mohamad Sudarsoino	92.85	80	92.33
16	Mohammad Rosi Amin	91.8	66	90.95
17	Muhammad Lutfi	88.79	65	90.85
18	Muhammad Subli	91.8	70	91.35
19	Riskiatal Hasanah	88.79	70	91.35
20	Rofiatul Ardawiyah	91.8	70	91.35
21	Ruly Rifky Hadi	90.85	75	91.84
22	Salim Umar	90.85	70	91.35
23	Sayid Abdullah	92.85	72	91.54
24	Siti Nur'aini	90.85	70	91.35
25	Siti Rofikah	90.85	75	91.84
26	Sri Wahyuni	90.85	70	91.35
27	Sulistiyowati	92.85	75	91.84
28	Sumiyati	91.8	70	91.35
29	Sunaryo	91.8	70	91.35
30	Tanwirotul Af'ideh	90.85	70	91.35

**4.4 Hasil K-means 5 Centroid
dan Regresi Linier
Sederhana Bahasa
Indonesia**

Centroid I : 6 Centroid II
 : 13 Centroid III : 17
 Centroid IV : 21 Centroid
 V : 26

Tabel 4.7K-means Iterasi4

No	Nama	Kluster1	Kluster2	Kluster3	Kluster4	Kluster5	Posisi
1	Agustina Fadilatul f.	4.88	5.62	7.31	3.93	1.94	5
2	Asanah	2.13	5.13	5.76	1.23	0.81	5
3	Ayyul Farikoinih	6.38	1.42	4.14	6.18	5.10	2
4	Fiqih Firmansyah	6.13	1.00	3.72	5.98	5.04	2
5	Hafiyatul Hasanah	1.88	5.16	5.67	0.99	1.06	4
6	Ilham Rahmat	0.38	5.89	5.29	1.40	3.31	1
7	Imam Syafi'i	1.80	5.17	5.64	0.92	1.14	4
8	Imron Rosidi	1.55	5.21	5.56	0.70	1.39	4
9	Imrotul Hasanah	3.30	5.17	6.32	2.36	0.36	5
10	Indah Purnama Sari	1.30	5.26	5.48	0.52	1.64	4
11	Matlubatul Hasanah	2.97	5.13	6.14	2.04	0.03	5
12	Misbahul Arifin	7.00	3.27	1.71	7.48	7.64	3
13	Moch. Wakil	5.02	2.09	0.69	5.42	5.58	3
14	Mochamad Yahya F.	5.55	5.92	7.78	4.60	2.61	5
15	Mohamad Sudarsono	2.05	5.14	5.73	1.15	0.89	5
16	Mohammad Rosi Amin	5.47	0.35	2.40	5.54	5.05	2
17	Muhammad Lutfi	5.06	3.34	0.68	5.68	6.23	3
18	Muhammad Subli	5.43	0.44	2.31	5.52	5.07	2
19	Riskiatal Hasanah	5.21	4.01	1.32	5.92	6.65	3
20	Rofiatul Ardawiyah	5.11	1.51	1.25	5.40	5.35	3
21	Ruly Rifky Hadi	1.13	5.31	5.44	0.43	1.81	4
22	Salim Umar	0.38	5.55	5.31	0.71	2.56	1
23	Sayid Abdullah	5.16	10.19	10.39	4.61	5.26	4
24	Siti Nur'aini	0.80	5.40	5.37	0.43	2.14	4
25	Siti Rofikah	0.72	5.43	5.36	0.47	2.22	4
26	Sri Wahyuni	1.55	5.21	5.56	0.70	1.39	4
27	Sulistiyowati	3.05	5.14	6.19	2.12	0.11	5
28	Sumiyati	5.73	0.27	2.98	5.70	5.00	2
29	Sunaryo	5.02	2.09	0.69	5.42	5.58	3
30	Tanwirotul Af'ideh	0.80	5.40	5.37	0.43	2.14	4

Tabel 4.8 Regresi Linier
 Sederhana

No	Nama	Nilai Cluster	Uas	Y Predikasi
1	Agustina Fadilatul f.	92.85	75	91.75
2	Asanah	92.85	70	91.28
3	Ayyul Farikoinih	92.8	72	91.47
4	Fiqih Firmansyah	92.8	65	90.82
5	Hafiyatul Hasanah	90.85	80	92.21
6	Ilham Rahmat	89.85	65	90.82
7	Imam Syafi'i	90.85	70	91.28
8	Imron Rosidi	90.85	65	90.82
9	Imrotul Hasanah	92.85	70	91.28
10	Indah Purnama Sari	90.85	75	91.75
11	Matlubatul Hasanah	92.85	70	91.28
12	Misbahul Arifin	89.8	66	90.91
13	Moch. Wakil	89.8	65	90.82
14	Mochamad Yahya F.	92.85	80	92.21
15	Mohamad Sudarsono	92.85	80	92.21
16	Mohammad Rosi Amin	92.8	66	90.91
17	Muhammad Lutfi	89.8	65	90.82
18	Muhammad Subli	92.8	70	91.28
19	Riskiatal Hasanah	89.8	70	91.28
20	Rofiatul Ardawiyah	89.8	70	91.28
21	Ruly Rifky Hadi	90.85	75	91.75
22	Salim Umar	89.85	70	91.28
23	Sayid Abdullah	90.85	72	91.47
24	Siti Nur'aini	90.85	70	91.28
25	Siti Rofikah	90.85	75	91.75
26	Sri Wahyuni	90.85	70	91.28
27	Sulistiyowati	92.85	75	91.75
28	Sumiyati	92.8	70	91.28
29	Sunaryo	89.8	70	91.28
30	Tanwirotul Af'ideh	90.85	70	91.28

4.5 Membandingkan Hasil Prediksi dengan UNAS

Tabel 4.9 Bahasa Indonesia

Ci (1), Cii (2), Ciii (3), CiiV (4), CV (5), CVI (6), CVII (7), CVIII (8), CIX (10), CX (12)										
No	Nama	JUMLAH CENTROID								Unas
		3	4	5	6	7	8	9	10	
1	Agustina Fadilatul f.	92.19	92.19	91.73	92.04	91.72	91.99	91.99	91.85	8.60
2	Asanah	91.56	91.56	91.21	91.31	91.25	91.44	91.27	91.14	7.40
3	Ayyul Farikoinih	91.81	91.81	91.42	91.60	91.44	91.66	91.56	91.43	6.20
4	Fiqih Firmansyah	90.92	90.92	90.68	90.57	90.79	90.89	90.55	90.44	7.00
5	Hafiyatul Hasanah	92.82	92.82	92.25	92.77	92.18	92.54	92.72	92.56	8.20
6	Ilham Rahmat	90.92	90.92	90.68	90.57	90.79	90.89	90.55	90.44	6.60
7	Imam Syafi'i	91.56	91.56	91.21	91.31	91.25	91.44	91.27	91.14	7.80
8	Imron Rosidi	90.92	90.92	90.68	90.57	90.79	90.89	90.55	90.44	6.00
9	Imrotul Hasanah	91.56	91.56	91.21	91.31	91.25	91.44	91.27	91.14	6.80
10	Indah Purnama Sari	92.19	92.19	91.73	92.04	91.72	91.99	91.99	91.85	8.60
11	Matlubatul Hasanah	91.56	91.56	91.21	91.31	91.25	91.44	91.27	91.14	7.80
12	Misbahul Arifin	91.05	91.05	90.79	90.72	90.88	91.00	90.70	90.58	4.60
13	Moch. Wakil	90.92	90.92	90.68	90.57	90.79	90.89	90.55	90.44	4.80
14	Mochamad Yahya F.	92.82	92.82	92.25	92.77	92.18	92.54	92.72	92.56	8.20
15	Mohamad Sudarsoino	92.82	92.82	92.25	92.77	92.18	92.54	92.72	92.56	8.40
16	Mohammad Rosi Amin	91.05	91.05	90.79	90.72	90.88	91.00	90.70	90.58	7.40
17	Muhammad Lutfi	90.92	90.92	90.68	90.57	90.79	90.89	90.55	90.44	7.40
18	Muhammad Subli	91.56	91.56	91.21	91.31	91.25	91.44	91.27	91.14	6.20
19	Riskiatal Hasanah	91.56	91.56	91.21	91.31	91.25	91.44	91.27	91.14	6.60
20	Rofiatul Ardawiyah	91.56	91.56	91.21	91.31	91.25	91.44	91.27	91.14	6.40
21	Ruly Rifky Hadi	92.19	92.19	91.73	92.04	91.72	91.99	91.99	91.85	8.00
22	Salim Umar	91.56	91.56	91.21	91.31	91.25	91.44	91.27	91.14	6.60
23	Sayid Abdullah	91.81	91.81	91.42	91.60	91.44	91.66	91.56	91.43	7.20
24	Siti Nur'aini	91.56	91.56	91.21	91.31	91.25	91.44	91.27	91.14	7.40
25	Siti Rofikah	92.19	92.19	91.73	92.04	91.72	91.99	91.99	91.85	7.80
26	Sri Wahyuni	91.56	91.56	91.21	91.31	91.25	91.44	91.27	91.14	7.80
27	Sulistiyowati	92.19	92.19	91.73	92.04	91.72	91.99	91.99	91.85	6.80
28	Sumiyati	91.56	91.56	91.21	91.31	91.25	91.44	91.27	91.14	7.00
29	Sunaryo	91.56	91.56	91.21	91.31	91.25	91.44	91.27	91.14	6.60
30	Tanwirotul Af'ideh	91.56	91.56	91.21	91.31	91.25	91.44	91.27	91.14	7.40
	RMSE	43.3	47.5	45.5	48.3	41.5	44.4	46.6	45.7	

Tabel 4.10 IPA

Ci (1), Cii (2), Ciii (3), CiiV (4), CV (5), CVI (6), CVII (7), CVIII (8), CIX (10), CX (12)										
No	Nama	JUMLAH								Unas
		3	4	5	6	7	8	9	10	
1	Agustina Fadilatul f.	76.3	76.4	76.1	77.1	77	76.5	77.3	76.7	6.00
2	Asanah	76.5	75.3	77.6	76.1	75	75.4	75.6	75.8	2.25
3	Ayyul Farikoinih	76.5	75.6	77.1	76.4	75.6	75.7	76.1	76.1	4.25
4	Fiqih Firmansyah	76.3	76.8	75.7	77.4	77.6	76.9	77.8	76.9	3.50
5	Hafiyatul Hasanah	76.3	76.6	75.8	77.3	77.4	76.8	77.6	76.8	5.25
6	Ilham Rahmat	76.5	75.3	77.6	76.1	75	75.4	75.6	75.8	3.75
7	Imam Syafi'i	76.4	76.3	76.3	77	76.8	76.4	77.1	76.6	5.00
8	Imron Rosidi	76.3	76.5	76	77.2	77.2	76.6	77.4	76.8	4.00
9	Imrotul Hasanah	76.4	76.2	76.4	76.9	76.6	76.3	76.9	76.5	4.75
10	Indah Purnama Sari	76.4	76.2	76.4	76.9	76.6	76.3	76.9	76.5	6.25
11	Matlubatul Hasanah	76.4	76	76.7	76.7	76.2	76.1	76.6	76.3	4.75
12	Misbahul Arifin	76.5	75.4	77.4	76.2	75.2	75.5	75.7	75.9	3.75
13	Moch. Wakil	76.5	75.4	77.4	76.2	75.2	75.5	75.7	75.9	4.00
14	Mochamad Yahya F.	76.4	76.2	76.4	76.9	76.6	76.3	76.9	76.5	6.00
15	Mohamad Sudarsoino	76.3	76.9	75.5	77.5	77.8	77	77.9	77	6.00
16	Mohammad Rosi Amin	76.5	75.6	77.1	76.4	75.6	75.7	76.1	76.1	3.75
17	Muhammad Lutfi	76.5	75.3	77.6	76.1	75	75.4	75.6	75.8	4.75
18	Muhammad Subli	76.4	75.7	77	76.5	75.8	75.8	76.2	76.1	5.00
19	Riskiatal Hasanah	76.4	76.1	76.6	76.8	76.4	76.2	76.7	76.4	3.50
20	Rofiatul Ardawiyah	76.5	75.5	77.3	76.3	75.4	75.6	75.9	76	4.00
21	Ruly Rifky Hadi	76.4	75.7	77	76.5	75.8	75.8	76.2	76.1	4.25
22	Salim Umar	76.4	75.7	77	76.5	75.8	75.8	76.2	76.1	3.50
23	Sayid Abdullah	76.5	75.6	77.1	76.4	75.6	75.7	76.1	76.1	4.75
24	Siti Nur'aini	76.6	74.8	78.2	75.7	74.2	74.9	74.9	75.4	3.50
25	Siti Rofikah	76.4	75.7	77	76.5	75.8	75.8	76.2	76.1	4.75
26	Sri Wahyuni	76.3	76.5	76	77.2	77.2	76.6	77.4	76.8	4.25
27	Sulistiyowati	76.3	76.8	75.7	77.4	77.6	76.9	77.8	76.9	4.25
28	Sumiyati	76.4	76.2	76.4	76.9	76.6	76.3	76.9	76.5	4.50
29	Sunaryo	76.5	75.4	77.4	76.2	75.2	75.5	75.7	75.9	3.00
30	Tanwirotul Af'ideh	76.4	76	76.7	76.7	76.2	76.1	76.6	76.3	5.00
	RMSE	22.3	17.4	21.3	25.5	13.8	19.1	20.7	23.3	

Tabel 4.11 Bahasa Inggris

CI (1), CII (2), CIII (3), CIV (4), CV (5), CVI (6), CVII (7), CVIII (8), CIX (9), CX (10)										
No	Nama	JUMLAH CENTROID								Unas
		3	4	5	6	7	8	9	10	
1	Agustina Fadilatul f.	82.1	82	81.8	81.6	82	81.6	82.1	82	6.20
2	Asanah	80.2	80	80	79.8	80	79.9	80.2	80.1	3.20
3	Ayyul Farikoinih	81	80.8	80.7	80.5	80.8	80.6	81	80.8	5.40
4	Fiqih Firmansysah	78.4	78	78.2	77.9	78.1	78.1	78.3	78.2	4.20
5	Hafiyatul Hasanah	83.9	84	83.6	83.5	83.9	83.4	84	83.9	5.60
6	Ilham Rahmat	78.4	78	78.2	77.9	78.1	78.1	78.3	78.2	3.40
7	Imam Syafi'i	80.2	80	80	79.8	80	79.9	80.2	80.1	3.60
8	Imron Rosidi	78.4	78	78.2	77.9	78.1	78.1	78.3	78.2	3.20
9	Imrotul Hasanah	80.2	80	80	79.8	80	79.9	80.2	80.1	4.40
10	Indah Purnama Sari	82.1	82	81.8	81.6	82	81.6	82.1	82	3.60
11	Matlubatul Hasanah	80.2	80	80	79.8	80	79.9	80.2	80.1	4.60
12	Misbahul Arifin	78.7	78.4	78.5	78.3	78.5	78.4	78.7	78.5	3.40
13	Moch. Wakil	78.4	78	78.2	77.9	78.1	78.1	78.3	78.2	4.60
14	Mochamad Yahya F.	83.9	84	83.6	83.5	83.9	83.4	84	83.9	6.40
15	Mohamad Sudarsoino	83.9	84	83.6	83.5	83.9	83.4	84	83.9	5.40
16	Mohammad Rosi Amin	78.7	78.4	78.5	78.3	78.5	78.4	78.7	78.5	3.80
17	Muhammad Lutfi	78.4	78	78.2	77.9	78.1	78.1	78.3	78.2	2.20
18	Muhammad Subli	80.2	80	80	79.8	80	79.9	80.2	80.1	3.20
19	Riskiatal Hasanah	80.2	80	80	79.8	80	79.9	80.2	80.1	3.40
20	Rofiatul Ardawiyah	80.2	80	80	79.8	80	79.9	80.2	80.1	3.20
21	Ruly Rifky Hadi	82.1	82	81.8	81.6	82	81.6	82.1	82	5.20
22	Salim Umar	80.2	80	80	79.8	80	79.9	80.2	80.1	3.20
23	Sayid Abdullah	81	80.8	80.7	80.5	80.8	80.6	81	80.8	4.80
24	Siti Nur'aini	80.2	80	80	79.8	80	79.9	80.2	80.1	4.60
25	Siti Rofikah	82.1	82	81.8	81.6	82	81.6	82.1	82	4.60
26	Sri Wahyuni	80.2	80	80	79.8	80	79.9	80.2	80.1	4.60
27	Sulistiyowati	82.1	82	81.8	81.6	82	81.6	82.1	82	4.60
28	Sumiyati	80.2	80	80	79.8	80	79.9	80.2	80.1	4.20
29	Sunaryo	80.2	80	80	79.8	80	79.9	80.2	80.1	3.40
30	Tanwirotul Af'ideh	80.2	80	80	79.8	80	79.9	80.2	80.1	4.60
	RMSE	6.8	7.7	5.3	6.6	8.2	6.4	7.5	6.9	

Tabel 4.12 Matematika

C1 (1), CII (2), CIII (3), CIV (4), CV (5), CVI (6), CVII (7), CVIII (8), CIX (9), CX (10)										
No	Nama	JUMLAH CENTROID								Unas
		3	4	5	6	7	8	9	10	
1	Agustina Fadilatul f.	82.7	84.6	86.4	86.4	86.6	86.2	86.3	85.1	6.75
2	Asanah	76.3	74.5	74	74.1	74.3	74.5	74.8	74.2	3.00
3	Ayyul Farikoinih	76.3	74.5	74	74.1	74.3	74.5	74.8	74.2	3.50
4	Fiqih Firmansyah	79.5	79.6	80.2	80.3	80.5	80.3	80.5	79.7	3.75
5	Hafiyatul Hasanah	77.5	76.5	76.4	76.6	76.8	76.8	77.1	76.4	2.75
6	Ilham Rahmat	76.3	74.5	74	74.1	74.3	74.5	74.8	74.2	3.75
7	Imam Syafi'i	78.2	77.5	77.7	77.8	78	78	78.2	77.5	2.75
8	Imron Rosidi	79.5	79.6	80.2	80.3	80.5	80.3	80.5	79.7	3.50
9	Imrotul Hasanah	76.3	74.5	74	74.1	74.3	74.5	74.8	74.2	4.00
10	Indah Purnama Sari	78.8	78.5	78.9	79	79.2	79.2	79.4	78.6	2.50
11	Matlubatul Hasanah	78.8	78.5	78.9	79	79.2	79.2	79.4	78.6	3.75
12	Misbahul Arifin	77.5	76.5	76.4	76.6	76.8	76.8	77.1	76.4	3.25
13	Moch. Wakil	78.8	78.5	78.9	79	79.2	79.2	79.4	78.6	3.25
14	Mochamad Yahya F.	80.8	81.6	82.7	82.7	82.9	82.7	82.8	81.9	7.25
15	Mohamad Sudarsaino	84	86.6	88.9	88.9	89.1	88.5	88.6	87.3	8.25
16	Mohammad Rosi Amin	76.3	74.5	74	74.1	74.3	74.5	74.8	74.2	3.25
17	Muhammad Lutfi	76.3	74.5	74	74.1	74.3	74.5	74.8	74.2	4.50
18	Muhammad Subli	76.3	74.5	74	74.1	74.3	74.5	74.8	74.2	4.00
19	Riskiatal Hasanah	76.3	74.5	74	74.1	74.3	74.5	74.8	74.2	2.50
20	Rofiatul Ardawiyah	76.9	75.5	75.2	75.3	75.5	75.7	76	75.3	2.75
21	Ruly Rifky Hadi	78.2	77.5	77.7	77.8	78	78	78.2	77.5	4.50
22	Salim Umar	76.3	74.5	74	74.1	74.3	74.5	74.8	74.2	3.00
23	Sayid Abdullah	76.3	74.5	74	74.1	74.3	74.5	74.8	74.2	3.75
24	Siti Nur'aini	78.8	78.5	78.9	79	79.2	79.2	79.4	78.6	4.25
25	Siti Rofikah	78.2	77.5	77.7	77.8	78	78	78.2	77.5	5.50
26	Sri Wahyuni	83.4	85.6	87.7	87.7	87.8	87.3	87.4	86.2	4.75
27	Sulistiyowati	78.8	78.5	78.9	79	79.2	79.2	79.4	78.6	2.75
28	Sumiyati	77.5	76.5	76.4	76.6	76.8	76.8	77.1	76.4	3.50
29	Sunaryo	76.3	74.5	74	74.1	74.3	74.5	74.8	74.2	2.75
30	Tanwirotul Af'ideh	76.3	74.5	74	74.1	74.3	74.5	74.8	74.2	3.00
	RMSE	48.9	53.6	49.4	47.3	51.2	44.8	57.8	44.3	

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang telah diuraikan pada bab-bab sebelumnya dapat diambil kesimpulan sebagai berikut :

1. Setelah menggunakan metode K-Means dengan beberapa centroid dan Regresi Linier Sederhana didapatkan hasil terbaik prosentase tingkat akurasi prediksi nilai UNAS siswa dengan metode RMSE (Root Means Squared Error) pada mata pelajaran bahasa Indonesia adalah 41.5 (7 Centroid), bahasa Inggris adalah 5.3 (5 Centroid), IPA adalah 13.8 (7 Centroid), dan Matematika adalah 44.3 (10 Centroid)
2. Didapatkan nilai prediksi UNAS mata pelajaran bahasa Inggris merupakan prediksi terbaik dengan rata-rata kesalahan terkecil 5.3

5.2 Saran

Adapun saran yang ditujukan untuk memberikan masukan yang lebih baik yaitu :

1. Penelitian yang akan datang bisa dikembangkan dengan menambah beberapa variabel supaya diperoleh hasil yang lebih baik dari penelitian ini.

DAFTAR PUSTAKA

- Muliadinata, Saban. “*Pengertian K - Means*” . Diakses pada 05 September 2014. <http://sharewy.blogspot.com/2013/04/algortma-k-means-clustering.html>.
- Margaretta., (2010). “*Pengertian Data Mining*”. Diakses pada 19 September 2014, <http://garethdata.blogspot.com/2010/03/pengertian-datamining.html>.
- Muhammad., (2013). “*Pengertian dan persamaan Regresi Linier*”. Diperoleh 20 September 2014, dari <http://jangkrik2011.blogspot.com/2013/05/regresi-linear-sederhana.html>.
- Cici., (2012). “*Penjelasan dan rumus RMSE (ROOT MEAN SQUARE ERROR) / Metode Rata – Rata Bergerak*”. Diperoleh 20 September 2014, dari http://directory.umm.ac.id/Data%20Elmu/doc/Cici_04610221.doc.
- Fadli., (2009). “*Analisis Regresi Linier Berganda*”. Diakses pada 23 September 2014, <http://dawaisimfoni.wordpress.com/karya-tulis-ilmiah2/metodologi-penelitian/analisis-regresi-2.pdf/>
- Yusuf, A., Ginardi, H., Arieshanti, I., “*Pengembangan Perangkat Lunak Prediktor Nilai Mahasiswa Menggunakan Metode Spectral Clustering dan Bagging Regresi Linier*”. Diakses pada 29 September 2014, dari <http://digilib.its.ac.id/public/ITS-paper-25685-5108100153-Paper.pdf/>