

**SISTEM INFORMASI PEMETAAN DISTRIBUSI GURU  
MENGUNAKAN *K-MEANS CLUSTERING*  
DAN METODE *ELBOW* GUNA MENDUKUNG KEBIJAKAN  
ZONASI PENDIDIKAN DI INDONESIA**

**Tesis  
untuk memenuhi sebagian persyaratan  
mencapai derajat Sarjana S-2 Program Studi  
Magister Sistem Informasi**



**Edy Umargono  
30000318410009**

**SEKOLAH PASCASARJANA  
UNIVERSITAS DIPONEGORO  
SEMARANG  
TAHUN 2020**

HALAMAN PENGESAHAN

TESIS

**SISTEM INFORMASI PEMETAAN DISTRIBUSI GURU MENGGUNAKAN  
K-MEANS CLUSTERING DAN METODE ELBOW GUNA Mendukung  
KEBIJAKAN ZONASI PENDIDIKAN DI INDONESIA**

Oleh:  
**Edy Umargono**  
30000318410009

Telah diujikan dan dinyatakan lulus ujian tesis pada tanggal 9 Juni 2020 oleh tim penguji Program Studi Magister Sistem Informasi Sekolah Pascasarjana Universitas Diponegoro.

Semarang, Juni 2020  
Mengetahui,

**Penguji I**



Drs. Bayu Surarso, M.Sc., Ph.D  
NIP. 196311051988031001

**Penguji II**



Dinar Mutiara Kusumo Nugraheni, S.T.,  
M.InfoTech.(Comp), Ph.D  
NIP. 197601102009122002

**Pembimbing I**



Jatmiko Endro Suseno, M.Si., Ph.D  
NIP. 197211211998021001

**Pembimbing II**



Vincencius Gunawan S.K., M.Si., Ph.D  
NIP. 197105221997021001

**Mengetahui :**  
**Dekan Sekolah Pascasarjana  
Universitas Diponegoro**

Dr. R.B. Sularto, S.H., M.Hum  
NIP. 196701011991031005

**Ketua Program Studi  
Magister Sistem Informasi**



Dr. Saryono, S.Si., M.Si  
NIP. 197306301998021001

**PERNYATAAN PERSETUJUAN  
PUBLIKASI TESIS UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Diponegoro, saya yang bertanda tangan di bawah ini :

Nama : Edy Umargono  
NIM : 30000318410009  
Program Studi : Magister Sistem Informasi  
Program : Sekolah Pascasarjana  
Jenis Karya : Tesis

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Diponegoro Hak Bebas Royalti Noneksklusif atas karya ilmiah saya yang berjudul :

**SISTEM INFORMASI PEMETAAN DISTRIBUSI GURU  
MENGUNAKAN *K-MEANS CLUSTERING* DAN METODE *ELBOW*  
GUNA Mendukung Kebijakan Zonasi Pendidikan  
DI INDONESIA**

beserta perangkat yang ada. Dengan Hak bebas Royalti Noneksklusif ini Program Studi Magister Sistem Informasi Sekolah Pascasarjana Universitas Diponegoro berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*) merawat, dan mempublikasikan tesis saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta.

Dibuat di : Semarang  
Pada tanggal : 21 Mei 2020  
Yang menyatakan



  
Edy Umargono  
NIM. 30000318410009

## PERNYATAAN

Dengan ini saya menyatakan bahwa dalam tesis ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar akademik di suatu perguruan tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Semarang, 21 Mei 2020



Edy Umargono

## KATA PENGANTAR

Alhamdulillah, puji syukur penulis panjatkan kehadirat Allah SWT, atas berkah, rahmat, hidayah dan karunia-Nya, sehingga tesis dengan judul **“SISTEM INFORMASI PEMETAAN DISTRIBUSI GURU MENGGUNAKAN *K-MEANS CLUSTERING* DAN METODE *ELBOW* GUNA MENDUKUNG KEBIJAKAN ZONASI PENDIDIKAN DI INDONESIA”** ini dapat diselesaikan. Tesis ini disusun untuk memenuhi salah satu persyaratan memperoleh gelar Magister Sistem Informasi Sekolah Pascasarjana Universitas Diponegoro Semarang.

Penyusunan tesis ini tidak lepas dari bantuan berbagai pihak baik secara langsung maupun tidak langsung, karena itu perkenankan penulis menyampaikan ucapan terima kasih kepada:

1. Bapak Jatmiko Endro Suseno, M.Si., Ph.D sebagai Pembimbing I dan Bapak Vincencius Gunawan S.K, M.Si., Ph.D sebagai Pembimbing II yang telah memberikan waktu untuk membimbing dan mengarahkan penulis dalam penyusunan tesis ini.
2. Bapak Dr. R.B. Sularto, S.H., M.Hum sebagai Dekan Sekolah Pascasarjana Universitas Diponegoro.
3. Bapak Dr. Suryono, M.Si sebagai Ketua Program Studi Magister Sistem Informasi Sekolah Pascasarjana Universitas Diponegoro.
4. Bapak Drs. Bayu Surarso, M.Sc., Ph.D dan Ibu Dinar Mutiara Kusumo Nugraheni, S.T., M.InfoTech.(Comp), Ph.D sebagai Penguji yang telah memberikan koreksi, masukan pada Tesis ini untuk menjadi lebih baik.
5. Bapak Drs. Wisnu Zaroh, M.Si, Kepala Badan Kepegawaian Daerah (BKD) Provinsi Jawa Tengah dan Bapak Mukhlison, S.Sos., M.M, Kepala Seksi Pendidik dan Tenaga Kependidikan SMA Dinas Pendidikan dan Kebudayaan Provinsi Jawa Tengah sebagai pengampu Instansi terkait yang telah memberikan masukan dan data yang dibutuhkan dalam tesis ini

Dengan keterbatasan pengalaman, pengetahuan maupun pustaka yang ditinjau, penulis menyadari bahwa tesis ini masih banyak kekurangan dan

pengembangan lebih lanjut masih diperlukan. Oleh sebab itu, penulis mengharapkan kritik dan saran sehingga tesis ini dapat dikembangkan dan bermanfaat untuk penelitian dan penulisan karya ilmiah khususnya di bidang sistem informasi di masa yang akan datang.

Semarang, 21 Mei 2020

Penulis,



Edy Umargono



## DAFTAR ISI

Halaman Judul.....	i
Halaman Pengesahan .....	<b>Error! Bookmark not defined.</b>
Halaman Pernyataan Persetujuan Publikasi.....	iii
Halaman Pernyataan.....	iv
Kata Pengantar .....	v
Daftar Isi.....	vii
Daftar Gambar.....	ix
Daftar Tabel .....	x
Daftar Lampiran .....	xi
Daftar Arti Lambang dan Singkatan .....	xii
Abstrak.....	xiii
Abstract .....	xiv
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Tujuan Penelitian.....	2
1.3. Manfaat Penelitian.....	2
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI.....	3
2.1 Tinjauan Pustaka.....	3
2.2. Dasar Teori .....	5
2.2.1. Aturan Kepegawaian Sumber Daya Aparatur Guru .....	5
2.2.2. <i>Data Mining</i> .....	8
2.2.3. <i>Clustering</i> .....	9
2.2.4. Algoritma <i>K-Means</i> .....	11
2.2.5. Metode <i>Elbow</i> .....	14
2.2.6. Validasi <i>Cluster</i> .....	16
BAB III METODE PENELITIAN.....	18
3.1 Bahan dan Alat Penelitian .....	18
3.2 Prosedur Penelitian .....	18

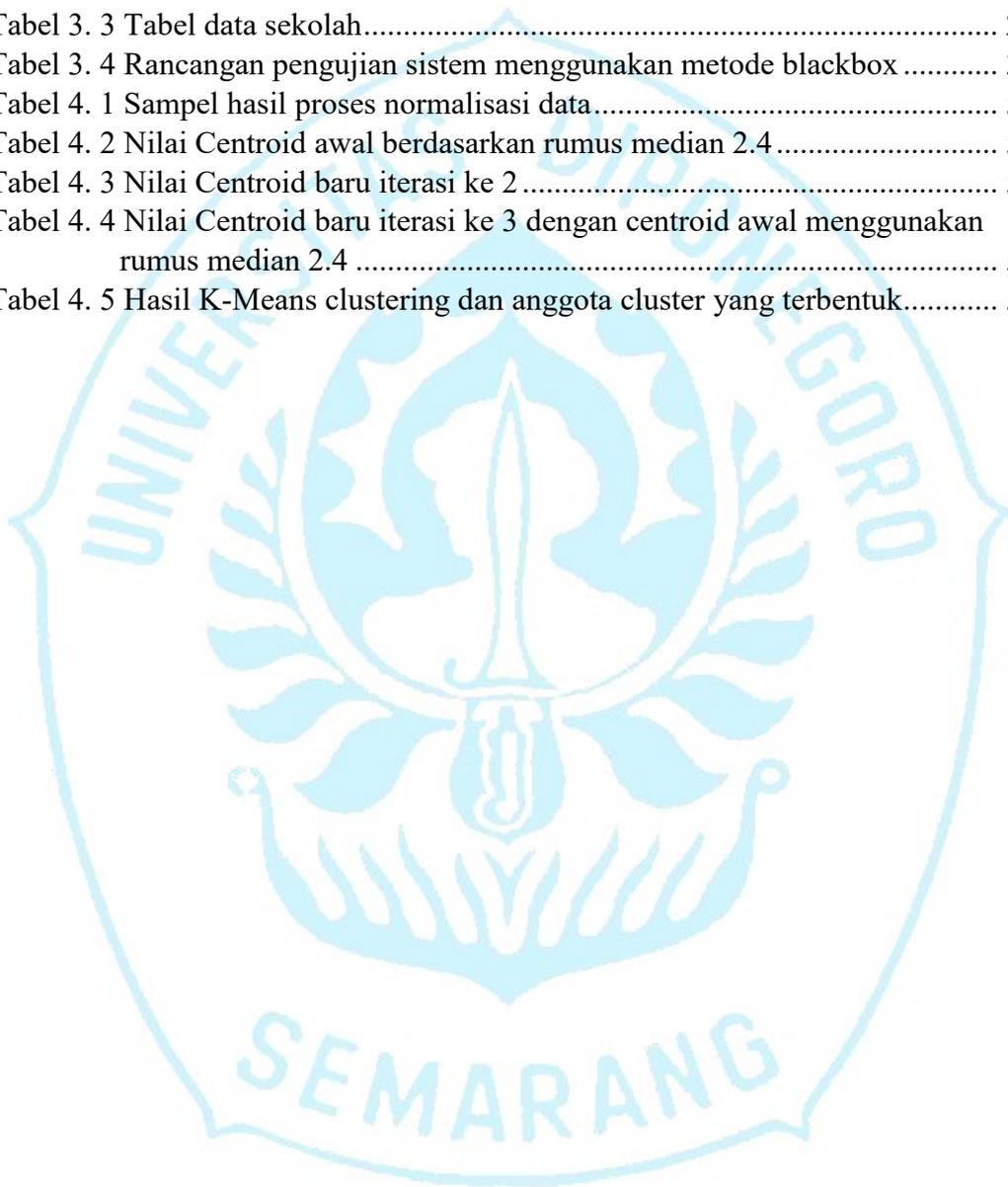
BAB IV HASIL PENELITIAN DAN PEMBAHASAN.....	41
4.1 Hasil penelitian.....	41
4.1.1. Implementasi Desain Antarmuka .....	41
4.1.2. Pengujian Sistem .....	50
4.2. Pembahasan .....	50
4.2.1. Analisa Data.....	50
4.2.2. Implementasi <i>K-Means Clustering</i> .....	52
4.2.3. Analisis Kinerja Sistem Berdasarkan Hasil Optimasi <i>Centroid</i> Awal .....	57
4.2.4. Optimasi Jumlah <i>Cluster</i> Dengan Menggunakan Metode <i>Elbow</i> .....	60
4.2.5. Interpretasi dan Validasi Anggota <i>Cluster</i> .....	62
BAB V KESIMPULAN DAN SARAN.....	66
5.1 Kesimpulan .....	66
5.2. Saran .....	67
DAFTAR PUSTAKA .....	68
DAFTAR LAMPIRAN.....	70

## DAFTAR GAMBAR

Gambar 2. 1 Grafik SSE terhadap jumlah cluster (Purnima dan Arvind, 2014)...	16
Gambar 3. 1 Prosedur penelitian.....	19
Gambar 3. 2 Kerangka sistem pemetaan Guru .....	22
Gambar 3. 3 Use case diagram sistem .....	25
Gambar 3. 4 Diagram Konteks sistem .....	24
Gambar 3. 5 Desain halaman login .....	29
Gambar 3. 6 Desain halaman utama .....	30
Gambar 3. 7 Input data sekolah .....	31
Gambar 3. 8 Desain tambah dan edit data manual.....	32
Gambar 3. 9 Desain halaman data awal .....	33
Gambar 3. 10 Desain halaman tampilan data normlisasi.....	33
Gambar 3. 11 Desain halaman metode Elbow .....	34
Gambar 3. 12 Desain pemilihan centroid awal dan jumlah cluster .....	35
Gambar 3. 13 Desain Detail Proses K-Means Clustering .....	36
Gambar 3. 14 Desain hasil akhir K-Means Clustering .....	36
Gambar 3. 15 Desain halaman visualisasi hasil akhir K-Means clustering pada peta online .....	37
Gambar 4. 1 Implementasi halaman login sistem .....	42
Gambar 4. 2 Implementasi halaman utama.....	42
Gambar 4. 3 Implementasi halaman input data sekolah .....	43
Gambar 4. 4 Implementasi input manual data pada sistem.....	44
Gambar 4. 5 Proses upload file berasal dari luar sistem .....	44
Gambar 4. 6 Fungsi edit data secara manual.....	44
Gambar 4. 7 Implementasi halaman data awal .....	45
Gambar 4. 8 Hasil normalisasi data .....	46
Gambar 4. 9 Implementasi metode Elbow dan hasilnya.....	46
Gambar 4. 10 Implementasi halaman penentuan centroid awal dan jml cluster...	47
Gambar 4. 11 Proses iterasi K-Means Clustering .....	48
Gambar 4. 12 Anggota hasil cluster yang terbentuk .....	48
Gambar 4. 13 Visualisasi anggota cluster yang terbentuk pada peta online.....	49
Gambar 4. 14 Hasil perbesaran visualisasi hasil cluster pada peta online .....	49
Gambar 4. 15 Statistik data jumlah siswa pada Sekolah Menengah Negeri.....	50
Gambar 4. 16 Statistik data jumlah guru, jumlah mata pelajaran dan jumlah rombongan belajar pada Sekolah Menengah Negeri.....	51
Gambar 4. 17 Hasil cluster percobaan K-Means clustering dengan menggunakan berbagai metode pemilihan centroid awal.....	59
Gambar 4. 18 Hasil perhitungan SSE pada aplikasi .....	61
Gambar 4. 19 Grafik Elbow .....	62

## DAFTAR TABEL

Tabel 3. 1 Gambaran umum Use Case.....	26
Tabel 3. 2 Tabel user.....	28
Tabel 3. 3 Tabel data sekolah.....	28
Tabel 3. 4 Rancangan pengujian sistem menggunakan metode blackbox .....	38
Tabel 4. 1 Sampel hasil proses normalisasi data.....	51
Tabel 4. 2 Nilai Centroid awal berdasarkan rumus median 2.4.....	53
Tabel 4. 3 Nilai Centroid baru iterasi ke 2 .....	56
Tabel 4. 4 Nilai Centroid baru iterasi ke 3 dengan centroid awal menggunakan rumus median 2.4 .....	56
Tabel 4. 5 Hasil K-Means clustering dan anggota cluster yang terbentuk.....	57



## DAFTAR LAMPIRAN

Lampiran 1. Narasi Use Case.....	70
Lampiran 2. Diagram Aktivitas.....	81
Lampiran 3. Diagram Sequence.....	92
Lampiran 4. Hasil pengujian sistem menggunakan metode Black Box Testing... 99	
Lampiran 5. Hasil SSE penentuan jumlah cluster optimal menggunakan metode Elbow .....	104
Lampiran 6. Nilai centroid awal dengan menggunakan berbagai rumus.....	117
Lampiran 7. Nilai centroid baru pada iterasi terakhir dengan menggunakan berbagai rumus.....	119
Lampiran 8. Hasil cluster berdasarkan nilai centroid awal menggunakan berbagai rumus .....	121
Lampiran 9. Contoh hasil aplikasi perhitungan iterasi awal s.d akhir 3 cluster..	124
lampiran 10. Nilai SSE Within, Nilai SSE Between, Nilai SSE Total 3 cluster.	128
Lampiran 11. Validasi anggota cluster dengan menggunakan form R10 .....	131
Lampiran 12. Sampel Form R10 dari Dinas Pendidikan dan Kebudayaan Pemerintah Provinsi Jawa Tengah.....	133
Lampiran 13. Visualisasi cluster hasil dari aplikasi Orange Data Mining.....	134

## DAFTAR ARTI LAMBANG DAN SINGKATAN

### DAFTAR ARTI LAMBANG

Lambang	Arti Lambang
$X^*$	Nilai hasil normalisasi data
$\mu_i$	Rumus mean sebagai centroid data
$d(x, c)$	Rumus jarak <i>Euclidean</i> dari objek data ke <i>centroid</i>
$x$	Nilai objek data
$j$	Jumlah Objek data
$RS$	Rumus heterogenitas <i>cluster</i>
$SS_T$	Rumus jumlah kuadrat total
$SS_w$	Rumus jumlah kuadrat dalam <i>cluster</i>
$ME$	Rumus median data

### DAFTAR SINGKATAN

Singkatan	Kepanjangan Singkatan
<i>STR</i>	<i>Student-Teacher Ratio</i>
<i>KDD</i>	<i>Knowledge Discovery From Data</i>
<i>SSE</i>	<i>Sum of Square Error</i>
<i>PHP</i>	<i>Hypertext Preprocessor</i>
<i>UML</i>	<i>Unified Modelling Language</i>
<i>CI</i>	<i>Code Igniter</i>
<i>MD5</i>	<i>Message-Digest Algoritihm 5</i>

**SISTEM INFORMASI PEMETAAN DISTRIBUSI GURU  
MENGUNAKAN *K-MEANS CLUSTERING* DAN METODE *ELBOW*  
GUNA Mendukung Kebijakan Zonasi Pendidikan  
DI INDONESIA**

**ABSTRAK**

Pemetaan distribusi Guru merupakan upaya mencapai salah satu tujuan kebijakan Zonasi Pendidikan di Indonesia yaitu pemerataan distribusi Guru pada sekolah-sekolah. Pemetaan dilakukan dengan mengelompokkan sekolah berdasarkan kebutuhan dan ketersediaan Guru, menggunakan metode *clustering* dengan algoritma *K-Means*. Algoritma ini memiliki kelemahan pada penentuan jumlah *cluster* berdasarkan asumsi dan sangat bergantung pada penentuan *centroid* awal. Untuk mengatasi kelemahan tersebut, pada penelitian ini digunakan metode *Elbow* untuk penentuan jumlah *cluster* optimal dan perbandingan *centroid* awal berdasarkan *mean* atau *median* dengan *centroid* awal secara acak. Hasil penelitian ini adalah Sistem Informasi Pemetaan distribusi Guru yang diujicobakan pada data Sekolah Menengah Atas Negeri (SMAN) di Kota Semarang, Kabupaten Semarang dan Kota Salatiga untuk mengetahui sekolah yang kelebihan atau kekurangan Guru kemudian dianalisis pengaruh penggunaan metode pada kinerja sistem. Hasil analisis kinerja sistem pada proses iterasi dengan jumlah *cluster* optimal menggunakan metode *Elbow* diperoleh efisiensi iterasi 22,86% dengan tingkat heterogenitas *cluster* sebesar 75,86%, sedangkan penggunaan *centroid* awal berdasarkan nilai *mean* atau *median* diperoleh efisiensi iterasi 25% dengan *cluster* yang konsisten dibanding dengan menggunakan *centroid* awal secara acak. Interpretasi hasil *cluster* yang divalidasi menggunakan Rekapitulasi Data Pendidik dan Tenaga Kependidikan (Form R10) menyatakan bahwa *cluster* satu dengan enam anggota *cluster* merupakan *cluster* dengan sekolah-sekolah yang kelebihan dan ketercukupan Guru, *cluster* dua dengan enam belas anggota *cluster* merupakan *cluster* dengan kekurangan Guru paling banyak dibandingkan dengan *cluster* tiga dengan delapan anggota *cluster*.

Kata kunci: metode partisi, *K-Means*, metode *Elbow*, optimasi *centroid*, pemetaan guru

# **TEACHER DISTRIBUTION MAPPING INFORMATION SYSTEM USING K-MEANS CLUSTERING AND ELBOW METHOD TO SUPPORT EDUCATION ZONATION POLICY IN INDONESIA**

## **ABSTRACT**

Mapping teacher distribution is an effort to achieve one of the objectives of the Education Zoning policy in Indonesia, namely equality of teacher distribution in schools. Mapping is done by grouping schools based on the needs and availability of teachers, using the clustering method with the K-Means algorithm. This algorithm has a weakness in determining the number of clusters based on assumptions and is very dependent on initial centroid determination. To overcome these weaknesses, in this study the Elbow method is used to determine the optimal number of clusters and the comparison of initial centroids based on the mean or median with the initial centroid randomly. The results of this study are the Teacher Distribution Mapping Information System which was tested on State High School (SMAN) data in Semarang City, Semarang Regency, and Salatiga City to find out which schools lack or excess teachers, then analyzed the effect of the use of methods on system performance. The results of the analysis of system performance in the iteration process with the optimal number of clusters using the Elbow method obtained an iteration efficiency of 22.86% with a cluster heterogeneity level of 75.86%, while the use of initial centroids based on the mean or median values obtained 25% iteration efficiency with consistent clusters compared by using random initial centroids. Interpretation of the results of the cluster which was validated using the Data Recapitulation of Educators and Education Personnel (Form R10) states that the cluster one with six cluster members is a cluster with schools with excess and sufficient teachers, cluster two with sixteen cluster members is the cluster with the most teachers lacking compared to cluster three with eight cluster members.

Keywords: partition method, K-Means, Elbow method, centroid optimization, teacher mapping.