

**SISTEM INFORMASI *CLUSTERING* KUALITAS PROGRAM STUDI
MENGUNAKAN OLAP (*ONLINE ANALYTICAL PROCESSING*) DAN K-MEDOIDS**

Tesis

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



Ghufron

30000317410010

**SEKOLAH PASCASARJANA
UNIVERSITAS DIPONEGORO
SEMARANG**

2020

HALAMAN PENGESAHAN

TESIS

**SISTEM INFORMASI *CLUSTERING* KUALITAS PROGRAM STUDI
MENGUNAKAN OLAP (*ONLINE ANALYTICAL PROCESSING*) DAN K-
MEDOIDS**

Oleh:
Ghufron
30000317410010

Telah diujikan dan dinyatakan lulus ujian tesis pada tanggal 8 Juni 2020 oleh tim penguji Program Studi Magister Sistem Informasi Sekolah Pascasarjana Universitas Diponegoro.

Semarang, 8 Juni 2020

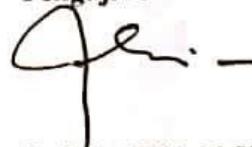
Mengetahui,

Penguji II

Penguji I

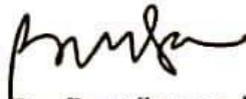


Dr. Suryono, S.Si., M.Si.
NIP. 197306301998021001



Farikhin, S.Si, M.Si, Ph.D.
NIP. 197312202000121001

Pembimbing I



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

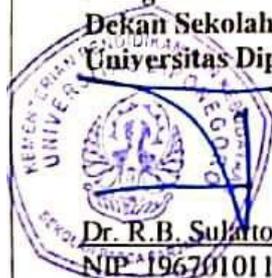
Pembimbing II



Dr. Rahmat Gemowo, M.Si.
NIP. 196511231994031003

Mengetahui:

**Dekan Sekolah Pascasarjana
Universitas Diponegoro**



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

**Ketua Program Studi
Magister Sistem Informasi**



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

**PERNYATAAN PERSETUJUAN
PUBLIKASI TESIS UNTUK KEPENTINGAN AKADEMIS**

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

Nama : Ghufron
NIM : 30000317410010
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 *Clustering* Kualitas Program Studi Menggunakan OLAP
(*Online Analytical Processing*) Dan K-medoids**

berserta 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 : 8 Juni 2020

Yang menyatakan


Ghufron

NIM. 30000317410010

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, 8 Juni 2020



Ghufron

KATA PENGANTAR

Puji dan syukur penulis ucapkan kepada Allah SWT, atas segala rahmat dan karunia-Nya yang telah memberikan nikmat kesehatan dan hikmat kepada penulis, sehingga penelitian ini dapat diselesaikan dengan baik sesuai dengan waktu yang telah direncanakan. Tesis berjudul “Sistem Informasi *Clustering* Kualitas Program Studi Menggunakan OLAP (*Online Analytical Processing*) Dan *K-medoids*”, disusun untuk memperoleh gelar Magister Komputer (M.Kom.) pada program studi Magister Sistem Informasi, Universitas Diponegoro. Penulis menyampaikan ucapan terima kasih setinggi-tingginya dan tak terhingga kepada yang terhormat :

1. Dr. R.B. Sularto, S.H., M.Hum., selaku Dekan Sekolah Pascasarjana Universitas Diponegoro Semarang.
2. Dr. Suryono, S.Si., M.Si selaku Ketua Program Studi Magister Sistem Informasi Sekolah Pascasarjana Universitas Diponegoro.
3. Drs. Bayu Surarso, M.Sc. Ph.D. selaku Pembimbing I terima kasih atas waktu, ilmu, saran, nasihat dan pengalaman yang bapak bagikan selama bimbingan.
4. Dr. Rahmat Gernowo, M.Si. selaku Pembimbing II yang penuh dengan kesabaran memberikan pengarahan dan banyak ilmu yang berguna dalam penulisan tesis ini.
5. Wakil Rektor I dan staff BAAk Universitas Pancasakti Tegal yang telah menyisihkan waktu dan membantu penulis dalam pengumpulan data penelitian.
6. Semua pihak yang tidak dapat disebutkan satu per-satu, telah membantu sampai dengan terselesaikannya tesis ini.

Penulis menyadari bahwa dalam penyusunan tesis masih jauh dari sempurna. Oleh karena itu, penulis mengharapkan saran dan kritik yang membangun. Penulis mohon maaf atas segala kekurangan dan kesalahan yang ada. Akhirnya, penulis berharap semoga tulisan ini bermanfaat. Aamiin.

Semarang, 8 Juni 2020

Penulis

DAFTAR ISI

	Halaman
Halaman Judul.....	i
Halaman Peengesahan	ii
Halaman Pernyataan Persetujuan Publikasi	iii
Kata Pengantar	v
Daftar Isi	vi
Daftar Gambar.....	vii
Daftar Tabel	ix
Daftar Lampiran	xi
Abstrak	xii
Abstrack	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan Penelitian	3
1.3 Manfaat Penelitian	3
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI.....	4
2.1 Tinjauan Pustaka	4
2.2 Data Warehouse	7
2.3 OLAP (Online Analytical Processing)	8
2.3.1 Keistimewaan OLAP	9
2.3.2 Teknik OLAP	9
2.3.3 Penyimpanan dalam OLAP.....	10
2.3.4 Desain Sistem Informasi	11
2.4 Metode Clustering K-Medoids.....	11
2.4.1 K-medoids	11
2.4.2 Data Pra-Pemrosesan	13
2.4.3 Fase Iteratif.....	15
2.4.4 Deteksi Outlier	16
2.4.5 Evaluasi Kemurnian Cluster	16
2.4.6 Algoritma Utama.....	17
2.5 Akreditasi Institusi Perguruan Tinggi	18
2.5.1 Matrik Akreditasi	19
2.6 Perhitungan Evaluasi clustering	21
2.7 Model Pengembangan Sistem Informasi	22
BAB III METODE PENELITIAN.....	23
3.1. Waktu dan Tempat	23
3.2. Bahan dan Alat Penelitian	23
3.3. Prosedur Penelitian.....	23
3.4. Data Flow Diagram	25
3.5. Perancangan Basis Data	27
3.6. Model Perancangan Skema Data Warehouse	30

3.7 Pemrosesan data dengan Algoritma K-medoids	32
BAB VI HASIL PENELITIAN DAN PEMBAHASAN	33
4.1 Hasil Penelitian	33
4.2 Analisis implementasi sistem	33
4.2.1 Pengumpulan Data	36
4.2.2 Data Ekstraksi	36
4.2.3 Data Transformasi	37
4.2.4 Data loading	39
4.2.5 Penyimpanan pada data warehouse	40
4.2.6 Model perancangan schema system	40
4.3 Desain Sistem	41
4.4 Implementasi Sistem	43
4.5 Pengujian sistem	60
4.5.1 Pengujian Operasi OLAP	60
4.5.1 Pengujian Clustering K-medoids	61
4.6 Proses Clustering K-medoid	61
4.6.1 Perhitungan data klaster K-medoids	65
BAB V KESIMPULAN DAN SARAN	71
5.1 Kesimpulan	71
5.2 Saran	71
DAFTAR PUSTAKA	72
LAMPIRAN	75

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Data Warehouse (Khalaf dan Obaid, 2017)	7
Gambar 2.2 Proses OLAP (Santoso dan Yulia, 2017)	8
Gambar 2.3 Flowchart Algoritma K-medoids	13
Gambar 2.4 Transformasi data (Dewi dkk., 2017)	14
Gambar 3.1 Diagram blok penelitian	24
Gambar 3.2 Kerangka Sistem Informasi	25
Gambar 3.3 diagram konteks level 0 sistem analisa data perguruan tinggi	26
Gambar 3.4 Data mentah Mahasiswa	27
Gambar 3.5 Data mentah penerimaan mahasiswa baru	27
Gambar 3.6 Data mentah prestasi mahasiswa	28
Gambar 3.7 Data tabel relasi dimensi olap sistem	29
Gambar 3.8 Model starflake pada tabel fakta mahasiswa	31
Gambar 4.1 Flowchart sistem OLAP	34
Gambar 4.2 Flowchart K-medoids clustering	35
Gambar 4.3 Data loading	39
Gambar 4.4 Data warehouse sistem	40
Gambar 4.5 Starflake schema OLAP	41
Gambar 4.6 Rancangan antarmuka sistem informasi	42
Gambar 4.7 Rancangan login sistem	42
Gambar 4.8 Halaman login	43
Gambar 4.9 Antarmuka Sistem Informasi Akreditasi Perguruan Tinggi	44
Gambar 4.10 Fitur utama migrasi data	44
Gambar 4.11 Administrasi pengguna	45
Gambar 4.12 Setting fitur hak akses pengguna	45
Gambar 4.13 Migrasi data	46
Gambar 4.14 Relasi antar tabel dimensi	46
Gambar 4.15 Hasil dimensi kuliah mahasiswa	47
Gambar 4.16 Data dimensi prestasi	47
Gambar 4.17 Data slicing OLAP dosen mengajar	48
Gambar 4.18 Data dimensi dosen mengajar	48
Gambar 4.19 Rata-rata ipk	49
Gambar 4.20 Grafik garis jumlah mahasiswa	49
Gambar 4.21 Rata-rata lulusan	50
Gambar 4.22 Dril down rata-rata program studi	50
Gambar 4.23 Data grafik mahasiswa program studi	51
Gambar 4.24 Data grafik calon mahasiswa program studi	51
Gambar 4.25 Grafik OLAP prestasi mahasiswa	52
Gambar 4.26 Grafik OLAP rata-rata IPK	52
Gambar 4.27 Grafik OLAP rata-rata lulusan	52
Gambar 4.28 Hasil analisis OLAP untuk clustering	53
Gambar 4.29 Data K-medoids clustering	54
Gambar 4.30 Data K-medoids clustering proses 1	55

Gambar 4.31 Data K-medoids clustering proses 2	55
Gambar 4.32 Data K-medoids clustering proses 3	56
Gambar 4.33 Data K-medoids clustering proses 3	56
Gambar 4.34 Hasil pengujian DBI.....	57
Gambar 4.35 Grafik pengujian DBI.....	58
Gambar 4.36 Grafik piechart jumlah program studi baik dan prodi cukup	58
Gambar 4.37 Program studi cluster cukup.....	58
Gambar 4.38 Program studi cluster baik.....	59

DAFTAR TABEL

	Halaman
Tabel 2.1 Kriteria Akreditasi.....	19
Tabel 2.2 Rasio jumlah calon mahasiswa lulus seleksi	19
Tabel 2.3 Lama studi mahasiswa dalam 3 tahun terakhir	20
Tabel 2.4 Rasio jumlah mahasiswa dan jumlah dosen tetap	20
Tabel 2.5 Rasio jumlah dosen tetap yang memiliki persyaratan dosen	20
Tabel 3.1 Struktur setting tahun akademik	30
Tabel 4.1 Data yang dikumpulkan untuk sumber data.....	36
Tabel 4.2 Data ekstraksi	37
Tabel 4.3 Data transformasi	37
Tabel 4.4 Hasil pengujian	60
Tabel 4.5 Jumlah Calon Mahasiswa.....	61
Tabel 4.6 Jumlah Mahasiswa	62
Tabel 4.7 Rata-rata masa studi mahasiswa	63
Tabel 4.8 Jumlah dosen pada program studi.....	64
Tabel 4.9 Jumlah dosen kualifikasi.....	65
Tabel 4.10 Data perguruan tinggi.....	66
Tabel 4.11 Iterasi ke 1 data perguruan tinggi.....	67
Tabel 4.12 iterasi ke 2 data perguruan tinggi	68
Tabel 4.13 Hasil pengujian davies bouldin index	70

DAFTAR LAMPIRAN

Halaman	
Lampiran 1. Surat Penelitian.....	75
Lampiran 2. Execute SQL Server Olap.....	76
Lampiran 3. Kode K-medoids Clustering	76
Lampiran 4. Multiple koneksi database cakephp.....	82

**SISTEM INFORMASI *CLUSTERING* KUALITAS PROGRAM STUDI
MENGUNAKAN OLAP (*ONLINE ANALYTICAL PROCESSING*) DAN K-
MEDOIDS**

ABSTRAKS

Penelitian tentang analisis data pada perguruan tinggi saat ini sangatlah penting dan memiliki banyak manfaat. Penelitian tersebut didasarkan belum adanya analisis data secara realtime *data warehouse* perguruan tinggi mengakibatkan informasi yang didapat kurang efektif. Oleh karena itu dibutuhkan sistem yang mampu melakukan kluster data kualitas program studi pada perguruan tinggi. *Online Analytical Processing* (OLAP) dan K-medoids *clustering* digunakan pada *clustering* kualitas program studi. OLAP sistem digunakan untuk menganalisis data dilakukan proses *Extract, Transform* dan *Loading* kemudian data dimasukkan kedalam *data warehouse* sedangkan K-medoids digunakan untuk kluster data kualitas program studi untuk mengelompokkan program studi masuk kategori baik maupun program studi cukup. Penelitian ini menghasilkan sistem informasi untuk melakukan analisa data *realtime* serta pengambilan keputusan strategis terhadap data akademik dan non akademik dan K-medoids *clustering* digunakan untuk mengkluster kualitas program studi termasuk kluster baik maupun program studi cukup untuk evaluasi perguruan tinggi kedepannya, dengan menggunakan pengujian evaluasi kluster *davies bouldin index* kluster data terbaik yaitu 0,4 yang dapat melakukan *clustering* kualitas program studi secara *realtime*.

Kata kunci: *clustering, data warehouse, OLAP, k-medoids, davies bouldin index.*

QUALITY CLUSTERING INFORMATION SYSTEM OF STUDY PROGRAM USING ONLINE ANALYTICAL PROCESSING AND K-MEDOIDS

ABSTRACT

Research on data analysis at tertiary institutions is now very important and has many benefits. The research is based on the absence of data analysis in realtime at the college data warehouse, resulting in the information obtained is less effective. Therefore we need a system that is able to cluster data quality in study programs at tertiary institutions. Online Analytical Processing (OLAP) and K-medoids clustering are used in clustering the quality of study programs. OLAP system is used to analyze the data in the process of Extract, Transform, and Loading then the data is entered into the data warehouse while K-medoids are used to cluster study program quality data to classify study programs into both good and sufficient study programs. This research resulted in an information system for analyzing real-time data and strategic decision making on academic and non-academic data and K-medoids clustering is used to cluster the quality of study programs including both good clusters and study programs for future college evaluations, by using cluster evaluation testing davies The best Bouldin cluster data index is 0.4 which can cluster the quality of study programs in real-time.

Keywords: clustering, data warehouse, OLAP, k-medoids, davies bouldin index.