

**KLASTERISASI PENDAFTARAN MAHASISWA  
BARU BERDASARKAN KOTA ASAL SEKOLAH  
DENGAN METODE K MEANS**

**SKRIPSI**

**Oleh :**

**CHOIRUL UMAM**

**201610225218**



**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS BHAYANGKARA JAKARTA RAYA  
2020**

## LEMBAR PERSETUJUAN PEMBIMBING

Judul Skripsi : Klasterisasi Pendaftaran Mahasiswa Baru  
Berdasarkan Kota Asal Sekolah Dengan  
Metode K - Means  
Nama Mahasiswa : Choirul Umam  
Nomor Pokok Mahasiswa : 201610225218  
Program Studi/Fakultas : Teknik Informatika/Teknik  
Tanggal Lulus Ujian Skripsi : Bekasi, 27 Juli 2020

Bekasi, 3 Agustus 2020

MENYETUJUI,

Pembimbing I

Pembimbing II



Prima Dina Atika, S.Kom., M.Kom.  
NIDN 0311037107



Sugiyatno, S.Kom., M.Kom.  
NIDN 0313077206

## LEMBAR PENGESAHAN

Judul Skripsi : Klasterisasi Pendaftaran Mahasiswa Baru  
Berdasarkan Kota Asal Sekolah Dengan  
Metode K – Means  
Nama Mahasiswa : Choirul Umam  
Nomor Pokok Mahasiswa : 201610225218  
Program Studi/Fakultas : Teknik Informatika/Teknik  
Tanggal Lulus Ujian Skripsi : Bekasi, 27 Juli 2020

Bekasi, 3 Agustus 2020

MENGESAHKAN,

Ketua Tim Penguji : Achmad Noe'man, S.Kom., M.Kom. .....  
NIDN : 0328048402

Penguji I : Siti Setiawati, S.Pd., M.Pd. .....  
NIDN : 0313107904


Penguji II : Prima Dina Atika, S.Kom., M.Kom. .....  
NIDN : 0311037107

MENGETAHUI,

Ketua Program Studi  
Teknik Informatika

  
Sugiyatno S.Kom., M.Kom.  
NIDN 0313077206

Dekan  
Teknik Informatika

  
Ismaniah S.Si., M.M.  
NIDN 0309036503

## LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa:

Skripsi yang berjudul Klasterisasi Pendaftaran Mahasiswa Baru Berdasarkan Kota Asal Sekolah Dengan Metode K - Means, ini adalah benar - benar merupakan hasil karya saya sendiri dan tidak mengandung materi yang ditulis oleh orang lain kecuali pengutipan sebagai referensi yang sumbernya telah dituliskan secara jelas sesuai dengan kaidah penulisan karya ilmiah.

Apabila di kemudian hari ditemukan adanya kecurangan dalam karya ini, saya bersedia menerima sanksi dari Universitas Bhayangkara Jakarta Raya sesuai dengan peraturan yang berlaku.

Saya mengizinkan skripsi ini dipinjam dan digandakan melalui Perpustakaan Universitas Bhayangkara Jakarta Raya.

Saya memberikan izin kepada Perpustakaan Universitas Bhayangkara Jakarta Raya untuk menyimpan skripsi ini dalam bentuk digital dan mempublikasikannya melalui internet selama publikasi tersebut melalui portal Universitas Bhayangkara Jakarta Raya.

Bekasi, 3 Agustus 2020

Yang membuat pernyataan,



Choirul Umam

201610225218

## ABSTRAK

**Choirul Umam. 201610225218.** *Klasterisasi pendaftaran mahasiswa baru berdasarkan kota asal sekolah dengan metode k – means.*

Skripsi ini membahas tentang proses pendaftaran mahasiswa baru Universitas X akan menghasilkan data yang berlimpah berupa data mahasiswa baru. Data mining memanfaatkan dan mengelola data – data yang terkumpul dari pendaftaran mahasiswa baru dari tahun 2016 sampai dengan 2019 di Universitas X. Penulisan ini bertujuan untuk menerapkan data mining dengan metode clustering k – means menggunakan excel dan rapidminer dan menentukan strategi promosi yang tepat pada cluster yang terbentuk. Metode dalam penulisan ini dilakukan berdasarkan tahapan Cross - Industry Standard Process - Data Mining (CRISP - DM). Atribut yang digunakan adalah id registrasi, kota asal sekolah, kategori sekolah, dan jenis sekolah. Hasil dari cluster mahasiswa yang terbentuk setelah proses clustering k – means adalah tiga klaster, dengan cluster pertama berjumlah 116 data, cluster kedua berjumlah 44 data, dan cluster ketiga berjumlah 142 data. Cluster dengan jumlah terbanyak adalah cluster tiga. Sehingga dapat digunakan untuk menentukan strategi promosi pendaftaran mahasiswa baru Universitas X.

**Kata kunci :** Klasterisasi Pendaftaran, Data Mining, CRISP-DM, K - Means

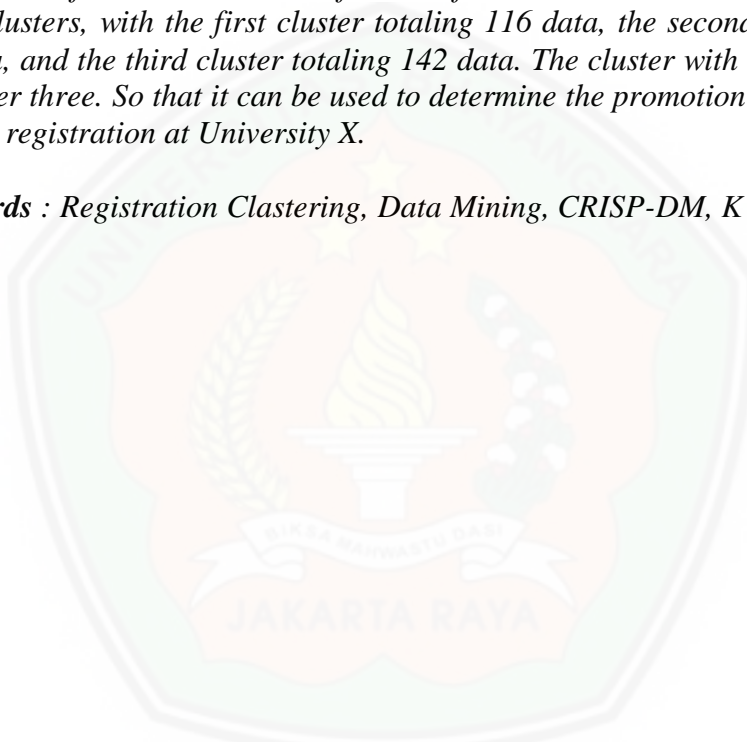


## ABSTRACT

**Choirul Umam. 201610225218.** *New student registration clustering based on the city of origin of the school using the k - means method.*

*This thesis discusses the process of registering new students of University X will produce abundant data in the form of new student data. Data mining utilizes and manages data collected from new student registrations from 2016 to 2019 at University X. This writing aims to implement data mining with the k-means clustering method using excel and rapidminer and determine the right promotion strategy in the cluster. formed. The method in this paper is based on the stages of Cross - Industry Standard Process - Data Mining (CRISP - DM). The attributes used are registration id, school's home town, school category, and type of school. The results of the student cluster formed after the k-means clustering process are three clusters, with the first cluster totaling 116 data, the second cluster totaling 44 data, and the third cluster totaling 142 data. The cluster with the most number is cluster three. So that it can be used to determine the promotion strategy for new student registration at University X.*

**Keywords :** *Registration Clustering, Data Mining, CRISP-DM, K - Means*



## LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : Choirul Umam  
NPM : 201610225218  
Program Studi : Teknik Informatika  
Fakultas : Teknik  
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya Hak Bebas Royalti Non- Eksklusif (*Non –Exclusive Royalty-Right*). Atas karya ilmiah saya yang berjudul :

### **“ Klasterisasi Pendaftaran Mahasiswa Baru Berdasarkan Kota Asal Sekolah Dengan Metode K - Means ”**

Beserta perangkat yang ada (bila diperlukan). Dengan hak yang bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalih media/formatkan, mengelolanya dalam bentuk basis data (*database*), mendistribusikan dan menampilkan/mempublikasikan di internet atau media lain untuk kepentingan akademis tanpa perlu permintaan ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat : Bekasi

Pada Tanggal : 3 Agustus 2020

Yang menyatakan,



Choirul Umam

## KATA PENGANTAR

Puji dan syukur ke hadirat Allah SWT, yang telah memberikan Rahmat, Hidayah dan Karunia-nya memperkenankan penulis menyelesaikan skripsi ini dengan judul “Klasterisasi pendaftaran mahasiswa baru berdasarkan kota asal sekolah dengan metode k – means”.

Penulis mengucapkan banyak terima kasih kepada kedua orang tua yang telah membesarkan dan mendidik penulis sebagai semangat dan motivasi terbesar dalam kehidupan anak – anaknya dan pihak – pihak telah membantu dalam menyelesaikan penulisan skripsi ini :

1. Irjen Pol (Purn) Drs. Bambang Karsono S.H., M.M. selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Ibu Ismaniah S.Si., M.M. selaku Dekan Fakultas Teknik Informatika Universitas Bhayangkara Jakarta Raya.
3. Bapak Sugiyatno S.Kom., M.Kom. selaku ketua Program Studi Teknik Informatika
4. Ibu Siti Setiawati, S.Pd., M.Pd. selaku dosen pembimbing akademik
5. Bapak Sugiyatno S.Kom., M.Kom. selaku dosen pembimbing skripsi
6. Ibu Prima Dina Atika S.Kom., M.Kom. selaku dosen pembimbing skripsi
7. Pak Achmad Noe'man, S.Kom., M.Kom. selaku dosen ketua tim penguji
8. Ibu Siti Setiawati, S.Pd., M.Pd. selaku dosen penguji I
9. Ibu Prima Dina Atika, S.Kom., M.Kom. selaku dosen penguji II
10. Kawan – Kawan Fakultas Teknik Universitas Bahayangkara Jakarta Raya Angkatan 2016/2017

Penulis menyadari bahwa mungkin masih terdapat banyak kekurangan dalam penulisan ini, untuk itu penulis mengharapkan kritik dan saran yang bersifat membangun, guna perbaikan di masa yang akan datang.



Akhir kata, semoga skripsi ini dapat bermanfaat bagi semua pihak yang telah meluangkan waktunya untuk membaca skripsi ini

Bekasi, 3 Agustus 2020



**Choirul Umam**



## DAFTAR ISI

	Halaman
<b>LEMBAR PERSETUJUAN PEMBIMBING .....</b>	<b>ii</b>
<b>LEMBAR PENGESAHAN.....</b>	<b>iii</b>
<b>LEMBAR PERNYATAAN .....</b>	<b>iv</b>
<b>ABSTRAK.....</b>	<b>v</b>
<b>ABSTRACT .....</b>	<b>vi</b>
<b>LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS.....</b>	<b>vii</b>
<b>KATA PENGANTAR.....</b>	<b>viii</b>
<b>DAFTAR ISI .....</b>	<b>x</b>
<b>DAFTAR TABEL .....</b>	<b>xiii</b>
<b>DAFTAR GAMBAR.....</b>	<b>xv</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>xvi</b>
<b>BAB I PENDAHULUAN .....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Identifikasi Masalah .....	3
1.3 Rumusan Masalah .....	3
1.4 Batasan Masalah.....	3
1.5 Tujuan dan Manfaat Penelitian .....	4
1.5.1 Tujuan Penelitian .....	4
1.5.2 Manfaat Penelitian .....	4
1.6 Tempat dan Waktu Penelitian.....	4
1.6.1 Tempat Penelitian .....	4
1.6.2 Waktu Penelitian.....	4
1.7 Metode Penelitian.....	5
1.8 Sistem Penulisan .....	5

<b>BAB II LANDASAN TEORI.....</b>	<b>7</b>
2.1 Tinjauan Pustaka .....	7
2.2 Landasan Teori.....	10
2.2.1 Definisi Data mining .....	10
2.2.2 Pengelompokan Teknik Data Mining .....	11
2.2.3 CRISP - DM .....	12
2.2.4 Clustering.....	13
2.2.5 K – MEANS Clustering .....	15
<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>17</b>
3.1 Objek Penelitian .....	17
3.2 Kerangka Pemikiran .....	17
3.3 Analisa Sistem Berjalan.....	18
3.3.1 Business Understanding .....	18
3.3.2 Data Understanding.....	18
3.4 Permasalahan .....	25
3.5 Analisis Usulan Sistem.....	25
3.5.1 Data Preparation.....	25
3.6 Analisis Kebutuhan Sistem.....	44
3.6.1 Kebutuhan Perangkat Keras .....	45
3.6.2 Kebutuhan Perangkat Lunak.....	45
<b>BAB IV PERANCANGAN SISTEM DAN IMPLEMENTASI.....</b>	<b>46</b>
4.1 Perancangan Sistem.....	46
4.1.1 Desain Model.....	46
4.2 Implementasi.....	48
4.2.1 Model Rapidminer .....	48
4.2.2 Evaluation.....	49

<b>BAB V PENUTUP</b> .....	52
5.1 Kesimpulan.....	52
5.2 Saran.....	52

**DAFTAR PUSTAKA**

**LAMPIRAN**

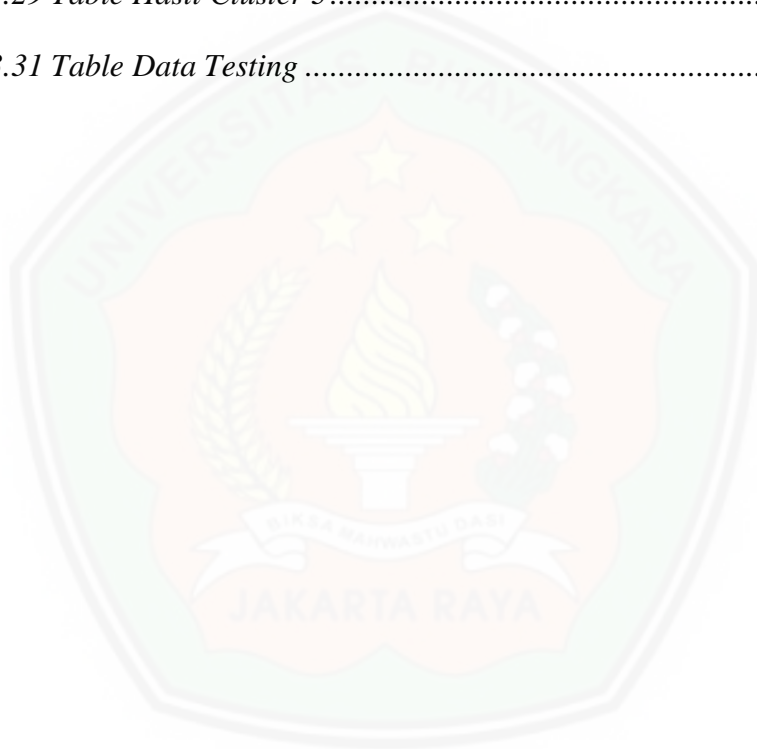




## DAFTAR TABEL

	Halaman
<i>Table 2.1 Table Tinjauan Pustaka</i> .....	8
<i>Table 3.1 Table Data 2016</i> .....	19
<i>Table 3.2 Table Data 2017</i> .....	20
<i>Table 3.3 Table Data 2018</i> .....	21
<i>Table 3.4 Table Data 2019</i> .....	23
<i>Table 3.5 Table Data Integration</i> .....	26
<i>Table 3.6 Table Cleaning Data</i> .....	27
<i>Table 3.7 Table Sampel dan Populasi</i> .....	28
<i>Table 3.8 Table Transformasi Kota Asal Sekolah</i> .....	29
<i>Table 3.9 Table Transformasi Kategori Sekolah</i> .....	31
<i>Table 3.10 Table Transformasi Jenis Sekolah</i> .....	31
<i>Table 3.11 Table Training Clustering K - Means</i> .....	32
<i>Table 3.12 Table Titik Pusat Awal Cluster</i> .....	33
<i>Table 3.13 Table literasi 1</i> .....	33
<i>Table 3.14 Cluster Baru Literasi 1</i> .....	34
<i>Table 3.15 Table Literasi 2</i> .....	35
<i>Table 3.16 Cluster Baru Literasi 2</i> .....	35
<i>Table 3.17 Table Literasi 3</i> .....	36
<i>Table 3.18 Cluster Baru Literasi 3</i> .....	36
<i>Table 3.19 Table Literasi 4</i> .....	37
<i>Table 3. 20 Cluster Baru Literasi 4</i> .....	37
<i>Table 3.21 Table Literasi 5</i> .....	38

<i>Table 3.22 Cluster Baru Literasi 5 .....</i>	<i>38</i>
<i>Table 3.23 Pengelompokan Pusat Cluster .....</i>	<i>39</i>
<i>Table 3.24 Table Titik Pusat Cluster Baru (Cluster 1).....</i>	<i>40</i>
<i>Table 3.25 Table Titik Pusat Cluster Baru (Cluster 2).....</i>	<i>40</i>
<i>Table 3.26 Table Titik Pusat Cluster Baru (Cluster 3).....</i>	<i>40</i>
<i>Table 3.27 Table Hasil Cluster 1.....</i>	<i>41</i>
<i>Table 3.28 Table Hasil Cluster 2.....</i>	<i>42</i>
<i>Table 3.29 Table Hasil Cluster 3.....</i>	<i>43</i>
<i>Table 3.31 Table Data Testing .....</i>	<i>43</i>



## DAFTAR GAMBAR

	Halaman
<i>Gambar 2.1 Gambar CRISP - DM .....</i>	12
<i>Gambar 3.1 Kerangka Pemikiran.....</i>	17
<i>Gambar 4.1 Gambar Model Normal.....</i>	46
<i>Gambar 4.2 Gambar Desain Model Split Data.....</i>	47
<i>Gambar 4.3 Gambar Rapidminer Data Normal.....</i>	48
<i>Gambar 4 4 Model Rapidminer Split Data .....</i>	48
<i>Gambar 4.5 Gambar Cluster Model Data Normal.....</i>	49
<i>Gambar 4.6 Gambar PerformanceVector Data Normal .....</i>	50
<i>Gambar 4.7 Gambar Cluster Model Split Data .....</i>	50
<i>Gambar 4.8 Gambar PerformanceVector Split Data .....</i>	51

## DAFTAR LAMPIRAN

1. Lampiran 1 Lembar Plagiasi
2. Lampiran 2 Biodata Mahasiswa
3. Lampiran 3 Kartu Bimbingan I
4. Lampiran 4 Kartu Bimbingan II

