

SKRIPSI

**SEGMENTASI MAHASISWA DENGAN ALGORITMA
UNSUPERVISED LEARNING DATA UNTUK
PREDIKSI PENERIMAAN MAHASISWA
BARU DI UNIVERSITAS
NAROTAMA**



DISUSUN OLEH :

ARIF ABRIYANTO

NIM :04315038

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS NAROTAMA

SURABAYA

2019

SKRIPSI

**SEGMENTASI MAHASISWA DENGAN ALGORITMA
UNSUPERVISED LEARNING DATA UNTUK
PREDIKSI PENERIMAAN MAHASISWA
BARU DI UNIVERSITAS
NAROTAMA**

Disusun Oleh:

ARIF ABRIYANTO

NIM: 04315038

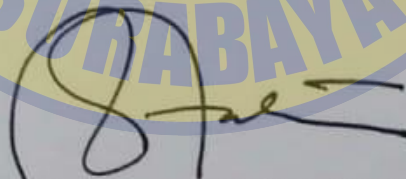
Diajukan guna memenuhi persyaratan
untuk memperoleh gelar Sarjana Komputer (S.Kom)

Pada Program Studi Sistem Informasi
Fakultas Ilmu Komputer
Universitas Narotama Surabaya

Surabaya, 26 Juli 2019

PRO PATRIA
menyetujui

Dosen Pembimbing



Natalia Damastuti, ST., M.T

NIDN: 0713047704

LEMBAR PENGESAHAN

SEGMENTASI MAHASISWA DENGAN ALGORITMA
UNSUPERVISED LEARNING DATA UNTUK
PREDIKSI PENERIMAAN MAHASISWA
BARU DI UNIVERSITAS
NAROTAMA

ARIF ABRIYANTO

NIM: 04315038

Dipertahankan di depan Penguji Skripsi
Program Studi Sistem Informasi
Fakultas Ilmu Komputer
Universitas Narotama Surabaya
Tanggal : 26 Juli 2019

Penguji,

Ketua Program Studi,

PRO PATRIA

1. Made Kamisutara, ST., M.Kom.
NIDN: 0706027501

Made Kamisutara, ST., M.Kom.
NIDN: 0706027501

2. Achmad Zakki Falani, S.Kom., M.Kom
NIDN: 0712058401

Fakultas Ilmu Komputer
Dekan,

3. Natalia Damastuti, ST., M.T
NIDN: 0713047704

Arvo Nugroho, ST., MT.
NIDN: 0721077001

SURAT PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Skripsi ini tidak terdapat karya yang pernah di ajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat Karya atau Pendapat yang pernah ditulis oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar acuan atau Daftar Pustaka.

Apabila ditemukan suatu Jiplakan/Plagiat maka saya bersedia menerima akibat berupa sanksi Akademis dan sanksi lain yang diberikan oleh yang berwenang sesuai dengan ketentuan peraturan dan perundang-undangan yang berlaku.

Surabaya, 26 Juli 2018



Arif Abriyanto
NIM. 04315038

MOTTO DAN PERSEMBAHAN

MOTTO

Jangan Pernah Menyerah Walau Banyak Rintangan Dan Cobaan Selih Berganti

“Karna Habis Hujan Terbitlah Pelangi”

PERSEMBAHAN

Puji syukur kepada Allah S.W.T atas limpahan rahmat dan kasihnya, hingga penulisan skripsi ini dapat saya selesaikan. Skripsi ini saya persembahkan kepada kedua orang tua terkasih, yang senantiasa membimbing, mendoakan, dan melimpahkan kasih sayang kepada putra-putri mereka. Segala bentuk perhatian dan nasehat dari kedua orang tua memberikan saya dorongan untuk melakukan yang terbaik dalam menyelesaikan permasalahan yang saya hadapi selama proses penyusunan tugas akhir ini. Tak lupa juga ucapan terima kasih kepada para sahabat, staf, dan dosen Universitas Narotama yang terlibat dalam pembuatan skripsi ini dari awal hingga akhir.

Saya menyadari bahwa Skripsi merupakan kewajiban yang harus dipenuhi sebagai syarat untuk mendapatkan gelar akademik. Akan tetapi lebih dari sebuah kewajiban, saya berharap Skripsi ini mampu memberikan manfaat kepada orang lain dan menjadi acuan serta referensi bagi para peneliti yang ingin melakukan penelitian dibidang yang sama. Dengan begitu penelitian yang saya lakukan tidak

selesai sampai di sini. Akhir kata, saya sangat bersyukur pengerjaan Skripsi ini dapat berjalan dengan lancar. Semoga semua yang saya kerjakan mendapatkan

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT yang telah melimpahkan rahmat dan hidayahnya sehingga penulis mampu menyelesaikan Skripsi ini, yang berjudul **“SEGMENTASI MAHASISWA DENGAN ALGORITMA UNSUPERVISED LEARNING DATA UNTUK PREDIKSI MAHASIWA BARU DI UNIVERSITAS NAROTAMA”**

Skripsi atau dikenal sebagai tugas akhir bagi mahasiswa dikerjakan demi memenuhi syarat akademis untuk menyelesaikan program Sarjana (S1) jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Narotama. Lebih dari itu hasil dari penulisan tugas akhir ini juga diharapkan dapat memberikan manfaat bagi orang lain, terutama para peneliti yang berminat pada penelitian dibidang yang serupa.

ABSTRAK

Semakin bertambahnya tahun teknologi informasi akan lebih muda untuk berkembang dengan cepat dan hampir mencakup disegala bidang kehidupan. Kemajuan ini dapat menghasilkan tersedianya data yang sangat besar dan banyak mulai dari bidang industri, ekonomi, dan pendidikan serta berbagai bidang kehidupan lainnya. Penerapan teknologi informasi dalam dunia pendidikan juga dapat menghasilkan banyak data yang berlimpah dari siswa yang mengikut prosesi pendidikan dan kegiatan belajar mengajar. Proses pendaftaran peserta didik baru atau mahasiswa baru Universitas Narotama menghasilkan data mahasiswa yang sangat berlimpah berupa data profil mahasiswa dan data kegiatan belajar mengajar. Hal ini pasti akan dilakukan berulang ulang sehingga menimbulkan penumpukan data mahasiswa, dan dapat mempengaruhi pencarian informasi terhadap data tersebut. Penelitian ini bertujuan untuk melakukan pengelompokan terhadap data mahasiswa Universitas Narotama dengan memanfaatkan proses data mining dengan menggunakan teknik Clustering. Metode yang digunakan untuk penelitian kali ini menggunakan *Kmeans Clustering*. K-Means merupakan salah satu metode data non-hierarchical clustering yang dapat mengelompokkan data mahasiswa ke dalam beberapa cluster berdasarkan kemiripan dari data tersebut, sehingga data mahasiswa yang memiliki karakteristik yang sama. Implementasi menggunakan digunakan untuk membantu menemukan nilai yang akurat. Atribut yang digunakan adalah wilayah asal pendaftar, program studi dan umur mahasiswa. Cluster mahasiswa yang terbentuk adalah tiga cluster, dengan cluster pertama 1112 mahasiswa, cluster kedua 825 mahasiswa dan cluster ketiga sejumlah 744 mahasiswa. Hasil dari penelitian ini digunakan sebagai salah satu referensi mengenai persebaran pendaftar yang mempunyai propek yang baik bagi Universitas dan sebagai acuan pihak maretng untuk menentukan strategi yang tepat.

Kata kunci: wilayah asal, program studi, umur, k-means clustering, Mahasiswa baru, strategi promosi

ABSTRACT

Increasingly years information technology will be easier to develop rapidly and almost cover all areas of life. This progress can result in the availability of very large data and many ranging from the fields of industry, economics, and education and various other fields of life. The application of information technology in the education world can also produce abundant data from students who attend educational processions and teaching and learning activities. The registration process for new students at Narotama University produces student data that is very abundant in the form of student profile data and data on teaching and learning activities. This will certainly be done repeatedly so as to cause a buildup of student data, and can affect the search for information on the data. This study aims to group the data of Narotama University students by utilizing the data mining process using Clustering techniques. This research method uses Kmeans Clustering. K-Means is one of the non-hierarchical clustering data methods that can group student data into several clusters based on the similarity of the data, thus student data has the same characteristics. Implementation is used to help find accurate values. The attributes used are the area of origin of the applicant, study program and age of the student. The student cluster formed was three clusters, the first cluster of 1112 students, the second cluster 825 students and the third cluster numbering 744 students. The results of this study are used as one of the references on the distribution of applicants who have good prospects for the University and as a reference for marketing to address the right strategy.

Keywords: origin area, study program, age, k-means clustering, new students, promotion strategies.

DAFTAR ISI

SKRIPSI	i
LEMBAR PENGESAHAN	ii
SURAT PERNYATAAN.....	iii
MOTTO DAN PERSEMBAHAN.....	iv
KATA PENGANTAR.....	v
ABSTRAK	vi
<i>ABSTRACT</i>	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR.....	x
DAFTAR TABEL	xi
PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Penulisan.....	4
BAB II.....	6
TINJUAN PUSTAKA	6
2.1 Tinjauan Penelitian Terdahulu	6
2.2 Teori Dasar Yang Digunakan.....	19
2.2.1 Strategi Promosi	19
2.2.2 Data mining.....	19
2.2.3 K-means.....	22
BAB III.....	24
METODOLOGI PENELITIAN	24
3.1 Tahapan Penelitian.....	24
3.2 identifikasi masalah.....	25
3.3 Studi Literatur	25
3.3 Pengumpulan Data.....	25

3.4 Prapemrosesan	26
3.5 Transformasi data	30
3.5 Evaluasi	34
3.6 Membuat Laporan.....	34
BAB IV	35
HASIL DAN PEMBAHASAN	35
4.1 Prapemrosesan Data	35
4.2 Data Mining.....	42
4.3 PERSEBARAN	47
BAB V	55
KESIMPULAN DAN SARAN	55
5.1 KESIMPULAN	55
5.2 SARAN	56
DAFTAR PUSTAKA	57



DAFTAR GAMBAR

Gambar 3.1 Tahapan penelitian	24
Gambar 3.2 pemodelan	27
Gambar 4.3.mengubah file XLS ke file CSV	38
Gambar 4.4Dilimiter Awal.....	39
Gambar 4.5 Mengubah Dilimiter	40
Gambar 4.6 Hasil Dilimiter yang telah diubah	41
Gambar 4.7 Proses menginputkan data.....	43
Gambar 4.8 memulai memproses.....	44
Gambar 4.9 Grafik hasil.....	46

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	10
Tabel 3.2 Inisialisasi wilayah	30
Tabel 3.3 Inisialisasi Progam Study	33
Tabel 3.4 Inisialisasi Umur	33
Tabel 4.5 data awal	36
Tabel 4.6 data yang akan digunakan	36
Tabel 4.7 centroid awal	45
Tabel 4.8 centroid akhir	45
Tabel 4.9 hasil cluster 1	47
Tabel 4.10 hasil cluster 2	49
Tabel 4.11 hasil cluster 3	51

