

**IMPLEMENTASI DATA MINING CLUSTERING MAHASISWA AKTIF  
ORGANISASI KEMAHASISWAAN UNIVERSITAS SATYA NEGARA  
INDONESIA MENGGUNAKAN ALGORITMA K-MEANS**

**Prabandini Kartika<sup>1</sup>, Safrizal<sup>2</sup>**

**Fakultas Teknik, Universitas Satya Negara Informatika  
Jl. Arteri Pondok Indah No. 11, Jakarta Selatan  
[kprabandini@gmail.com](mailto:kprabandini@gmail.com), [safrizal.st.mm@gmail.com](mailto:safrizal.st.mm@gmail.com)**

**ABSTRACT**

*One important aspect in evaluating the success of the work program at the University of Satya Negara Indonesia is the Student Organization of Universitas Satya Negara Indonesia (OK USNI). OK USNI allows students to have a study load taken each semester, activities that are followed, and academic activities can influence and test the quality of students who have an impact on the final GPA and the accuracy of graduating students. To prove that graduates who actively participate in OK USNI can increase their academic value and pass on time or not, a data grouping technique is used, namely K-Means Clustering. K-Means was chosen because it has a fairly high accuracy on the size of the object, so this algorithm is relatively more measured and efficient for processing large quantities of objects. The system development method uses the PHP programming language and MySQL for the database. The final results of this study in the form of grouping students produced 3 groups of OK USNI active graduates in C1 totaling 46 students or 64%, C2 totaling 22 students or 31%, and C3 totaling 4 students or 6%. For non-active USNI graduates who are in C1, there are 19 students or 26%, C2 is 45 students or 63%, and C3 is 8 students or 11%. This proves that by participating in the OK USNI, the duration of graduate studies is not affected.*

**Keywords:** *Students, Data Mining, Clustering, K-Means*

**ABSTRAK**

Salah satu aspek penting dalam evaluasi keberhasilan penyelenggaraan program kerja di Universitas Satya Negara Indonesia adalah Organisasi Kemahasiswaan Universitas Satya Negara Indonesia (OK USNI). OK USNI membuat mahasiswa memiliki beban studi yang diambil setiap semesternya, kegiatan yang diikuti, dan aktivitas akademik dapat mempengaruhi dan menguji kualitas dari mahasiswa yang berdampak pada IPK terakhir dan ketepatan lulus mahasiswa. Untuk membuktikan bahwa lulusan yang aktif mengikuti OK USNI dapat meningkat nilai akademiknya dan lulus tepat waktu atau tidak, maka digunakan suatu teknik pengelompokan data yaitu *K-Means Clustering*. *K-Means* dipilih karena memiliki ketelitian yang cukup tinggi terhadap ukuran objek, sehingga algoritma ini relatif lebih terukur dan efisien untuk pengolahan objek dalam jumlah besar. Metode pengembangan sistem menggunakan bahasa pemrograman PHP dan MySQL untuk *database*. Hasil akhir dari penelitian ini berupa pengelompokan mahasiswa dihasilkan 3 kelompok lulusan aktif OK USNI yang berada di C1 berjumlah 46 mahasiswa atau 64%, C2 berjumlah 22 mahasiswa atau 31%, dan C3 berjumlah 4 mahasiswa atau 6%. Bagi lulusan tidak aktif OK USNI yang berada di C1 berjumlah 19 mahasiswa atau 26%, C2 berjumlah 45 mahasiswa atau 63%, dan C3 berjumlah 8 mahasiswa atau 11%. Hal ini membuktikan bahwa dengan mengikuti OK USNI, lama studi lulusan tidak terpengaruh.

**Kata Kunci :** *Mahasiswa, Data Mining, Clustering, K-Means*

## PENDAHULUAN

### A. Latar Belakang

Universitas Satya Negara Indonesia (USNI) sudah memiliki pembinaan mahasiswa, baik itu dalam *hardskill* maupun *softskill*. Selain dalam bidang akademik, USNI juga mempunyai wadah lain untuk meningkatkan nilai dari setiap mahasiswanya melalui Organisasi Kemahasiswaan.

Organisasi Kemahasiswaan Universitas Satya Negara Indonesia (OK-USNI) merupakan salah satu aspek penting dalam evaluasi keberhasilan penyelenggaraan program kerja di Universitas Satya Negara Indonesia. Dengan adanya Organisasi Kemahasiswaan maka predikat Universitas juga akan semakin membaik dengan program-program kerja yang diajukan setiap Organisasi Kemahasiswaan. Beban studi yang diambil setiap semesternya, kegiatan yang diikuti, dan aktivitas akademik dapat mempengaruhi dan menguji kualitas dari mahasiswa yang berdampak pada IPK terakhir dan ketepatan lulus mahasiswa, dimana standar ketepatan lulus mahasiswa di USNI untuk program sarjana adalah empat tahun dan program diploma adalah tiga tahun. Untuk membuktikan bahwa lulusan yang aktif mengikuti Organisasi Kemahasiswaan dapat meningkatkan nilai akademiknya dan lulus tepat waktu atau tidak, maka digunakan suatu teknik pengelompokan data yaitu *clustering*.

Dari beberapa teknik *clustering* yang paling sederhana dan umum dikenal adalah *clustering* K-Means. K-Means merupakan salah satu metode data *clustering* non hirarki yang mempartisi data ke dalam cluster sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok lain.

Berdasarkan permasalahan tersebut, maka penulis berkeinginan untuk membuat suatu implementasi data mining *clustering* mahasiswa aktif Organisasi Kemahasiswaan Universitas Satya Negara Indonesia menggunakan Algoritma K-Means.

### B. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan sebelumnya, terdapat permasalahan yang akan dibahas dalam penelitian ini yaitu: “Bagaimana mengimplementasikan data mining *clustering* mahasiswa aktif organisasi kemahasiswaan Universitas Satya Negara Indonesia menggunakan metode algoritma *K-MEANS* ?”

### C. Batasan Masalah

Adapun batasan masalah yang ditunjukkan dalam penelitian ini adalah sebagai berikut :

1. Data yang dianalisis adalah data lulusan yang aktif di Organisasi Kemahasiswaan Universitas Satya Negara Indonesia jabatan 2013/2014 dan data lulusan yang tidak aktif tahun 2015, 2016, 2017.
2. Algoritma yang digunakan adalah algoritma K-Means *Clustering*.

### D. Tujuan dan Manfaat Penelitian

#### 1. Tujuan Penelitian

Tujuan dari penelitian ini adalah :

- a. Mengimplementasikan data mining untuk *clustering* mahasiswa aktif Organisasi Kemahasiswaan Universitas Satya Negara Indonesia.
- b. Menggunakan metode K-Means dalam pembuatan aplikasi *clustering* mahasiswa aktif Organisasi Kemahasiswaan Universitas Satya Negara Indonesia.

#### 2. Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah :

- a. Untuk mengetahui prestasi akademik mahasiswa baik lulusan yang mengikuti kegiatan Organisasi Kemahasiswaan maupun yang tidak mengikuti kegiatan Organisasi Kemahasiswaan.
- b. Untuk menambah pengetahuan dalam penggunaan algoritma *K-MEANS*.

## LANDASAN TEORI

### 1. Data Mining

*Data mining* merupakan proses otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial dan berguna yang bermanfaat yang tersimpan di dalam *database* besar (Turban et al : 2005). Pengelompokan data mining yaitu: Deskripsi, prediksi, klasifikasi, prediksi, pengklusteran, Asosiasi.

*Clustering* adalah proses pengelompokan kumpulan data menjadi beberapa kelompok sehingga objek di dalam satu kelompok memiliki banyak kesamaan dan memiliki banyak perbedaan dengan objek dikelompok lain (Han dan Kamber (2011),).

### 2. K-Means

Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster*/kelompok dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan *cluster*/kelompok yang lain sehingga data yang berada dalam satu *cluster*/kelompok memiliki tingkat variasi yang kecil (Agusta : 2007).

Secara historis, bentuk esensial *K-Means* ditemukan oleh sejumlah peneliti dari lintas disiplin ilmu. Yang paling berpengaruh adalah Llyod (1982), Forgey (1965), Friedman dan Rubin (1967), dan McQueen (1967). Algoritma *K-Means* berkembang hingga menjadi konteks yang lebih besar sebagai algoritma *hill-climbing*, seperti yang disampaikan oleh Gray dan Nuhoff (1998).

### 3. PHP

PHP (PHP: Hypertext Preprocessor) adalah Bahasa script yang ditanam di sisi server. Kalau kita pakai istilah sehari-hari, mungkin seperti ini: Prosesor PHP dijalankan di *server* (*Windows* atau *Linux*). Saat sebuah halaman dibuka dan mengandung kode PHP, prosesor itu akan menerjemahkan dan mengeksekusi semua perintah dalam halaman tersebut, dan kemudian menampilkan hasilnya ke *browser* sebagai halaman HTML biasa. (Prasetio, Adhi, 2012 : 122)

### 4. HTML (*Hyper Text Markup Language*)

HTML (*Hyper Text Markup Language*) adalah bahasa standar yang digunakan untuk pembuatan halaman *web* atau *world wide web*, dengan *hypertext* dan informasi lain yang akan ditampilkan pada halaman *web*. Dokumen *hypertext* bisa berisi teks, gambar, dan tipe informasi lain seperti data file, audio, video, dan program *executeable*. (Sutarman, 2012:163).

### 5. MySQL

MySQL adalah salah satu jenis *database server* yang sangat terkenal. Kepopulerannya disebabkan MySQL menggunakan SQL sebagai bahasa dasar untuk mengakses databasenya. Selain itu, ia bersifat *free* (Anda tidak perlu membayar untuk menggunakannya) pada berbagai platform (kecuali pada Windows, yang bersifat *shareware* atau Anda perlu membayar setelah melakukan evaluasi dan memutuskan untuk digunakan untuk keperluan produksi). (Kadir, Abdul, 2002 : 353)

### 6. Database

*Database* atau memiliki istilah lain Basis Data merupakan suatu kumpulan data yang saling berhubungan dan berkaitan dengan subjek tertentu pada tujuan tertentu pula. Hubungan antardata ini dapat dilihat oleh adanya *field* ataupun kolom. (Saputra, Agus, 2002 : 1)

### 7. Metode Pengembangan

Metodologi Pengembangan menggunakan Model waterfall yang terdiri dari : Analisa kebutuhan , Desain Sistem, Penulisan Koding program, pengujian Program, penerapan program dan pemeliharaan.

UML (*Unified Modeling Language*) adalah salah standar bahasa yang banyak digunakan di dunia industri untuk mendefinisikan *requirement*, membuat analisis & desain, serta menggambarkan arsitektur dalam pemrograman berorientasi objek. (Rosa et.al., 2013 : 133)

### 8. Universitas

Institut dan Universitas adalah institusi perguruan tinggi yang menyediakan pendidikan tinggi yang mengarah kepada level sarjana. Institut menawarkan pendidikan akademik dan/atau vokasi dalam kelompok disiplin ilmu pengetahuan, teknologi dan/atau seni tertentu. Disisi lain, Universitas menawarkan pendidikan akademik dan/atau vokasi dalam berbagai kelompok disiplin ilmu pengetahuan, teknologi dan/atau seni. Institusi pendidikan tinggi ini dapat juga melayani pendidikan pada level profesional. Institut Seni adalah salah satu contohnya. (DIKTI, 2011)

**9. Mahasiswa**

Mahasiswa adalah seseorang yang dalam proses menimba ilmu ataupun belajar dan terdaftar sedang menjalani pendidikan pada salah satu bentuk perguruan tinggi yang terdiri dari akademik, politeknik, sekolah tinggi, institute dan universitas. Ada beberapa status mahasiswa yaitu, mahasiswa aktif, mahasiswa non aktif, dan mahasiswa cuti akademik. Mahasiswa aktif adalah mahasiswa yang terdaftar pada semester tertentu sehingga berhak mengikuti kegiatan akademik serta mendapatkan layanan administratif dan akademik. Mahasiswa non aktif adalah mahasiswa yang tidak terdaftar pada semester tertentu tanpa izin Rektor. Mahasiswa cuti akademik adalah mahasiswa yang tidak terdaftar pada semester tertentu atas izin Rektor. (Hartaji, 2012 : 5)

**10. Organisasi**

Organisasi adalah suatu sistem perserikatan formal, berstruktur, dan terkoordinasi dari sekelompok orang yang bekerja sama dalam mencapai tujuan tertentu. Organisasi kemahasiswaan adalah wahana dan sarana pengembangan diri mahasiswa kearah perluasan wawasan dan pendekatan kecendekiawan serta integritas kepribadian manusia Pancasila yang cerdas, kritis, santun, bermoral, demokratis, bertanggung jawab, dan memiliki daya saing berdasarkan prinsip memanusiakan manusia sesuai dengan hakikat manusia. (Hasibuan, 2011:120)

## METODOLOGI PENELITIAN

### A. Waktu dan Tempat Penelitian

Waktu penelitian dilakukan selama 3 bulan, mulai April sampai bulan Juni 2018. Penelitian ini dilakukan di Universitas Satya Negara Indonesia yang berlokasi di Jl. Arteri Pondok Indah No. 11, Kebayoran Lama, Jakarta Selatan.

### B. Metode Pengumpulan Data

Adapun teknik pengumpulan data yang peneliti gunakan sebagai berikut:

1. Observasi  
Pada tahap ini peneliti memperoleh berbagai data dengan cara pengamatan dan peninjauan langsung terhadap objek penelitian untuk mengetahui gambaran yang terjadi pada Organisasi Kemahasiswaan Universitas Satya Negara Indonesia.
2. Wawancara  
Dengan cara mengajukan pertanyaan langsung kepada kepala Biro Administrasi Akademik Kemahasiswaan Perencanaan Sistem Informasi (BAAKPSI) untuk mendapatkan informasi yang lebih akurat.
3. Studi Pustaka  
Dengan cara mempelajari dan membaca literatur – literatur, catatan – catatan dan laporan – laporan yang ada hubungannya dengan permasalahan yang menjadi obyek penelitian.

### C. Tahapan – Tahapan Metode *K-Means*

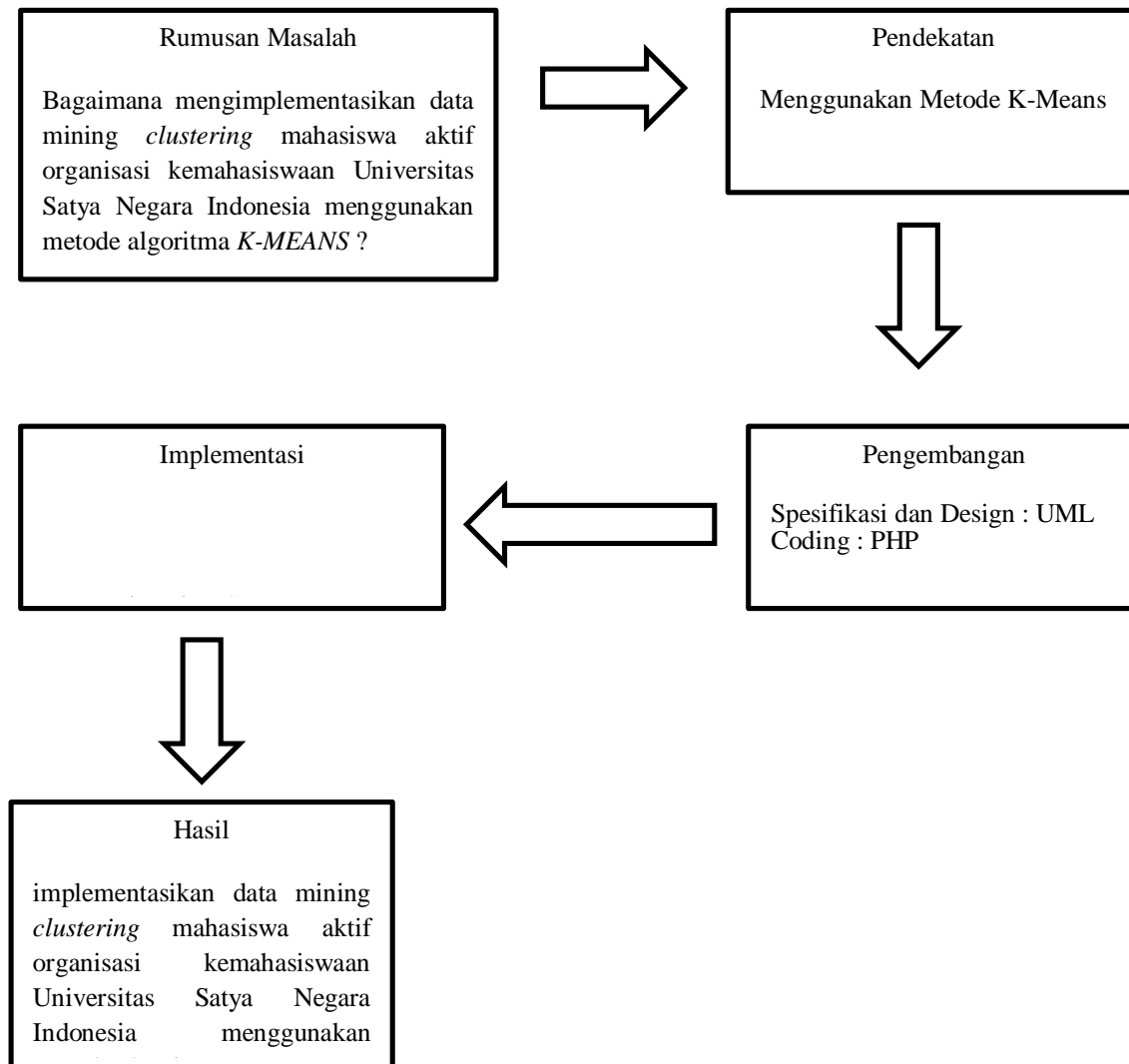
Secara garis besar, algoritma *K-Means Clustering* dijelaskan dalam 5 tahap berikut :

1. Inisialisasi: tentukan nilai *K* sebagai jumlah *cluster* yang diinginkan dan metrik ketidakmiripan (jarak) yang diinginkan. Jika perlu, tetapkan ambang batas perubahan fungsi objektif dan ambang batas perubahan posisi *centroid*, seperti dinyatakan dalam persamaan berikut:

$$d(x_1, c_1) = \sqrt{\sum_{i=1}^r (x_{1i} - c_{1i})^2}$$

2. Pilih *K* data dari set data *X* sebagai *centroid*.
3. Alokasikan semua data ke *centroid* terdekat dengan metrik jarak yang sudah ditetapkan (memperbarui *cluster ID* setiap data).
4. Hitung kembali *centroid* *C* berdasarkan data yang mengikuti cluster masing-masing.
5. Ulangi langkah 3 dan 4 hingga kondisi *konvergen* tercapai, yaitu (a) perubahan fungsi objektif sudah di bawah ambang batas yang diinginkan; atau (b) tidak ada data yang berpindah *cluster*; atau (c) perubahan posisi *centroid* sudah di bawah ambang batas yang ditetapkan.

#### D. Kerangka Berfikir



Gambar 1. Kerangka Pemikiran

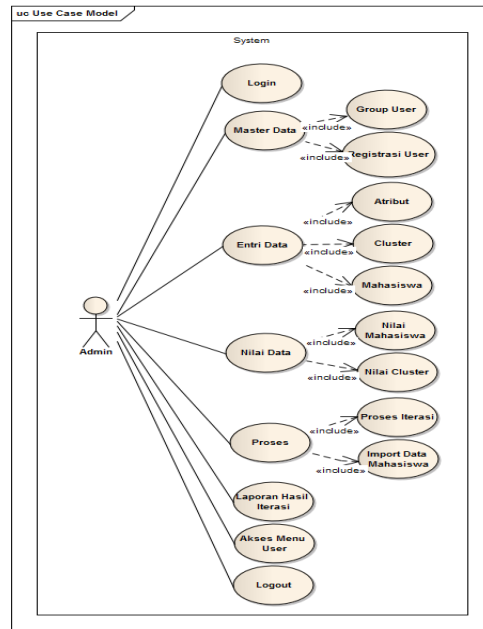
## 4. HASIL DAN PEMBAHASAN

### A. Perhitungan *K-Means*

Pada bagian ini akan dijelaskan mengenai penggunaan algoritma *k-means*. Dalam membentuk *clustering* (kelompok). Adapun langkah – langkah dalam perhitungan algoritma *k-means* seperti: inialisasi, menentukan pusat *centroid* awal, pencarian jarak terdekat terhadap *centroid*, dan pengelompokan jarak terdekat. Didapat data sebanyak 144 data mahasiswa, dengan Atribut yaitu: Nama Mahasiswa, Nilai IPK - dikategorikan dimensi 1, Lama Studi dikategorikan deimensi 2

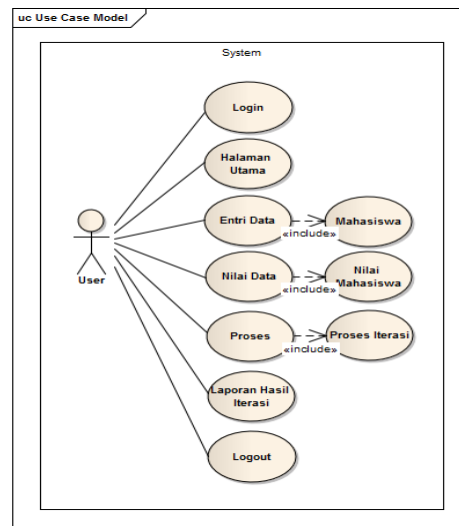
## B. Proses Perancangan Sistem

### Use Case Diagram



Gambar 2. Use Case Diagram Admin

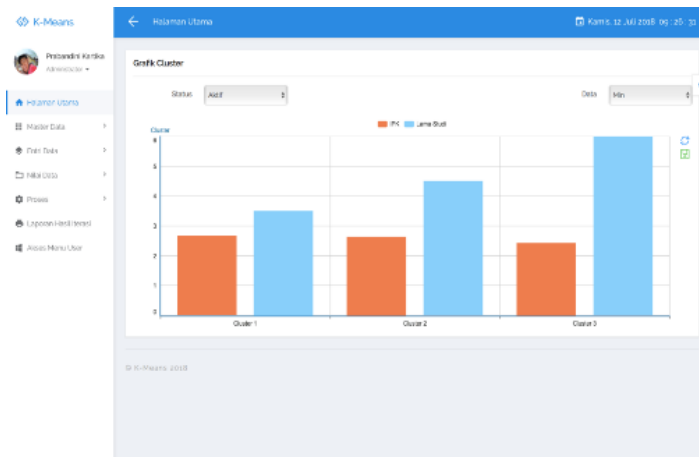
#### a. Use Case Diagram User



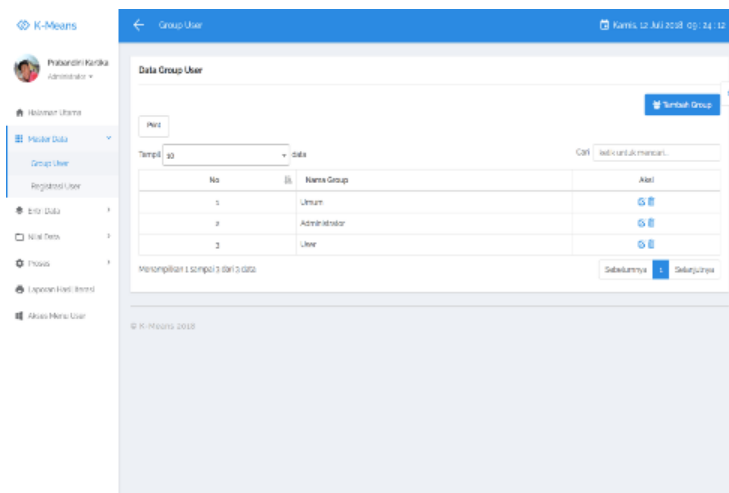
Gambar 3. Use Case Diagram User



## 1. Tampilan Halaman Utama



Gambar 4 Tampilan Halaman Utama



Gambar 5. Tampilan Menu Registrasi User

## KESIMPULAN

### A. Kesimpulan

Penelitian ini berfokus pada penerapan metode *clustering* terhadap analisis prestasi akademik bagi lulusan yang pernah mengikuti Organisasi kemahasiswaan. Kesimpulan yang diperoleh dari analisis yang telah dilakukan adalah dari klasterisasi lulusan yang aktif OK USNI C1 terdapat 46 data dan sekitar 64% dari 72 mahasiswa dengan IPK 2,66 – 3,54 dan lama studi 3,5 – 4 tahun, C2 terdapat 22 data dan sekitar 31% dari 72 mahasiswa dengan IPK 2,63 – 3,32 dan lama studi 4,5 – 5,5 tahun, dan C3 terdapat 4 data dan sekitar 6% dari 72 mahasiswa dengan IPK 2,42 – 2,88 dengan lama studi 6 – 7 tahun. Bagi lulusan yang tidak aktif Organisasi Kemahasiswaan diperoleh dari analisis yang telah dilakukan adalah

dari klusterisasi lulusan yang tidak aktif OK USNI C1 terdapat 19 data dan sekitar 26% dari 72 mahasiswa dengan IPK 2,77 – 3,78 dan lama studi 4 tahun, C2 terdapat 45 data dan dan sekitar 63% dari 72 mahasiswa dengan IPK 2,59 – 3,43 dan lama studi 4,5 – 5,5 tahun, dan C3 terdapat 8 data dan sekitar 11% dari 72 mahasiswa dengan IPK 2,58 – 3,14 dan lama studi 6 – 7 tahun. Hal ini membuktikan bahwa dengan mengikuti Organisasi Kemahasiswaan, lama studi lulusan tidak terpengaruh.

## B. Saran

Untuk penelitian selanjutnya, dapat menggunakan data lulusan yang lebih banyak, baik lulusan yang mengikuti Organisasi Kemahasiswaan maupun yang tidak mengikuti Organisasi Kemahasiswaan. Serta saran yang diberikan untuk penelitian aplikasi *K-MEANS* ini lebih lanjut adalah mencari nilai K yang terbaik karena sangat berpengaruh terhadap hasilnya nanti.

## Daftar Pustaka

- Agusta, Yudhi. 2007. 'K-Means penerapan permasalahan dan metode terkait'. Jurnal Sistem dan Informatika, Vol 3.
- Ayub, M., & Kristanti, T. (2013). Model Analisis Classification dan Clustering untuk Data Mahasiswa dan Dosen di Perguruan Tinggi. Bandung: Laporan Hibah Bersaing
- Berlin: Springer. Han, J., Kamber, M., & Pei, J. (2011). Data Mining Concepts and Techniques Third Edition. Waltham: Elsevier Inc.
- Cahyo Aji Nugroho, Clustering Kelompok Swadaya Masyarakat (KSM) Dalam Menentukan Kebijakan Bantuan Badan Pemberdayaan Masyarakat Di Kota Surabaya Dengan Menggunakan Metode Self-Organizing MAP (SOM) Dan K-Means, 2012
- Fina Nasari, Penerapan K-Means Clustering Pada Data Penerimaan Mahasiswa Baru (Studi Kasus : Universitas Potensi Utama), 2015, ISSN : 2302-3805
- Gorunescu, F. (2011). Data Mining Concepts, Models and Techniques.
- Lillyan Hadjaratie, Prediksi Dan Pemetaan Data Mahasiswa Fakultas Teknik Universitas Negeri Gorontalo Menggunakan Pendekatan Data Mining, 2015
- Narwati, Pengelompokan Mahasiswa Menggunakan Algoritma K-Means, 2015
- Rima Dias Ramadhani, Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Universitas Dian Nuswantoro, 2015
- Turban, E., dkk, 2005, Decicion Support Systems and Intelligent Systems, Andi Offse