

LAPORAN AKHIR
PENELITIAN DOSEN PEMULA



**PENGUNAAN *K-MEANS CLUSTERING METHODS* DALAM
PENGELOMPOKAN PENGGUNA INTERNET UNTUK
MENGOPTIMALKAN PENGGUNAAN INTERNET**

Tahun ke 1 dari rencana 1 tahun

TIM PENGUSUL

1. Ketua : Mukhtar Hanafi, ST., M.Cs., NIDN 0602047502
2. Anggota : Auliya Burhanuddin, S.Si., NIDN 0630058202

UNIVERSITAS MUHAMMADIYAH MAGELANG

Oktober 2017

HALAMAN PENGESAHAN

Judul : Penggunaan K-Means Clustering Methods dalam Pengelompokan Pengguna Internet untuk Mengoptimalkan Penggunaan Internet

Peneliti/Pelaksana
Nama Lengkap : MUKHTAR HANAFAI, S.T, M.Cs
Perguruan Tinggi : Universitas Muhammadiyah Magelang
NIDN : 0602047502
Jabatan Fungsional : Asisten Ahli
Program Studi : Teknik Informatika
Nomor HP : -
Alamat surel (e-mail) : hanafi@ummgl.ac.id

Anggota (1)
Nama Lengkap : AULIYA BURHANUDDIN S.Si
NIDN : 0630058202
Perguruan Tinggi : Universitas Muhammadiyah Magelang

Institusi Mitra (jika ada)
Nama Institusi Mitra : -
Alamat : -
Penanggung Jawab : -
Tahun Pelaksanaan : Tahun ke 1 dari rencana 1 tahun
Biaya Tahun Berjalan : Rp 19,000,000
Biaya Keseluruhan : Rp 19,000,000


Mengetahui,
Dekan Fakultas Teknik



(Yuni Anifah Fatimah, ST.,MT.,Ph.D)
NIP/NIK 987408139

Kab.Magelang, 31 - 10 - 2017

Ketua,



(MUKHTAR HANAFAI, S.T, M.Cs)
NIP/NIK 057508191

Menyetujui,
Ketua L.P3M



(Dr. Heny Setyowati ER, S Kp., M.Kes)
NIP/NIK 937008062

DAFTAR ISI

	Hlm
HALAMAN SAMBUNG	i.
HALAMAN PENGESAHAN.....	ii.
DAFTAR ISI.....	iii.
RINGKASAN	iv.
BAB 1 PENDAHULUAN	1
BAB 2 TINJAUAN PUSTAKA	3
BAB 3 TUJUAN DAN MANFAAT.....	9
BAB 4 METODE PENELITIAN.....	10
BAB 5 HASIL DAN PEMBAHASAN.....	14
BAB 6 KESIMPULAN.....	16
DAFTAR PUSTAKA	17
LAMPIRAN.....	18

RINGKASAN

Untuk mendukung aktivitas perkuliahan, Universitas Muhammadiyah Magelang (UM Magelang) memanfaatkan internet sebagai fasilitas untuk mendukung proses pembelajaran. Pembagian bandwidth internet yang dilakukan saat ini masih kurang optimal, karena ada perbedaan jumlah pengakses internet pada setiap fakultas. Hal lain yang mempengaruhi keoptimalan dari akses internet ini adalah tidak selalu jumlah pengakses internet banyak membutuhkan bandwidth yang besar, karena contain yang diakses banyak yang kecil, demikian juga sebaliknya, terkadang pengaksesnya sedikit tetapi bandwidth yang digunakan besar. Untuk mengatasi hal tersebut maka perlu dilakukan pembagian kuota bandwidth berdasarkan kelompok pengguna, agar pembagiannya dapat optimal sesuai dengan kelompok pengguna tersebut. Pada penelitian ini pembagian kelompok/*clustering* dilakukan dengan metode K Means. Metode K Means adalah metode pengelompokkan terawasi sehingga hasil kelompoknya dapat dijadikan rekomendasi untuk pembatasan kuota untuk tiap penggunanya. Untuk mengetahui kualitas hasil klasternya digunakan metode purity. Hasilnya, dari data 3000 pengguna internet aktif, direkomendasikan untuk dibagi menjadi 3, 4 atau 5 *cluster* pengguna, dengan nilai purity terbaiknya adalah 0.87 untuk 4 *cluster* pengguna internet.

Kata Kunci : *Trafik Internet, Clustering, K-Means, Purity*

BAB 1. PENDAHULUAN

1.1 Latar Belakang.

Pada saat ini internet sudah beralih menjadi kebutuhan yang wajib ada, terutama di suatu lembaga pendidikan. Dengan internet kita dapat terhubung dengan siapa saja tanpa terhalang oleh jarak. Karena internet adalah layanan jaringan dari komputer yang saling terhubung menggunakan sebuah perangkat jaringan agar bisa terkoneksi ke seluruh dunia. Internet mempunyai banyak manfaat terutama dalam mendukung proses belajar mengajar.

Universitas Muhammadiyah Magelang (UM Magelang) adalah salah satu perguruan tinggi di Magelang yang memiliki 6 fakultas dan 16 program studi dan pada tahun akademik 2015/2016 mempunyai mahasiswa aktif sebanyak 4777 mahasiswa. Untuk mendukung aktivitas perkuliahan, UM Magelang memanfaatkan internet sebagai fasilitas untuk mendukung proses pembelajaran. Bandwidth yang dimiliki dibagi rata ke tiap-tiap fakultas. Tetapi hal tersebut masih kurang optimal, karena ada perbedaan jumlah pengakses internet pada setiap fakultas. Pada waktu tertentu ada fakultas yang trafik internetnya sangat padat karena pengakses internet sangat banyak, pada waktu yang sama ada fakultas yang trafik internetnya sangat lengang karena pengaksesnya sedikit. Hal lain yang mempengaruhi keoptimalan dari akses internet ini adalah tidak selalu jumlah pengakses internet banyak membutuhkan bandwidth yang besar, karena contain yang diakses banyak yang kecil, demikian juga sebaliknya, terkadang pengaksesnya sedikit tetapi menghabiskan bandwidth yang besar. Untuk mengatasi hal tersebut maka perlu dibuat pembagian bandwidth dengan mengelompokkan penggunaanya agar pembagian bandwidthnya dapat dilakukan secara optimal.

Untuk mendapatkan hasil pengelompokan yang optimal, diperlukan analisis pengelompokan/*clustering* pada pengguna internet tersebut. Analisis cluster akan mengklasifikasi objek sehingga setiap objek yang paling dekat kesamaannya dengan objek lain yang berbeda dalam cluster yang sama. Kelompok-kelompok yang terbentuk memiliki homogenitas internal yang tinggi dan heterogenitas eksternal yang tinggi juga. Sedangkan metode yang digunakan dalam pengelompokan ini adalah dengan metode K-Means clustering. Pemilihan metode tersebut karena ada beberapa kelebihan pada algoritma K-Means, yaitu (S. Russel dan P. Norvig, 2010)

, yaitu: 1) Mudah untuk diimplementasikan dan dijalankan; 2) Waktu yang dibutuhkan untuk menjalankan pembelajaran ini relatif cepat; 3) Mudah untuk diadaptasi; dan 4) Umum digunakan. Selanjutnya dari hasil pengelompokan tersebut dapat menjadi rekomendasi atas besar kuota yang akan diberikan ketiap-tiap user id dalam satu semester.

1.2 Rumusan Masalah.

Rumusan masalah penelitian ini adalah apakah dengan menggunakan metode K-Means dapat diperoleh hasil pengelompokan yang optimal yang berpengaruh pada pembagian kuota bandwidth yang optimal juga.

BAB 2. TINJAUAN PUSTAKA

2.1. Teknologi WiFi

Perkembangan *WiFi* yang merupakan teknologi nirkabel telah menjadi suatu hal yang sangat populer dalam perkembangannya dalam jaringan komputer lokal (LAN). *WiFi* merupakan sekumpulan standar yang digunakan untuk Jaringan Lokal Nirkabel (*Wireless Local Area Network/WLAN*) berdasarkan pada spesifikasi IEEE 802.11. Rincian spesifikasi *WiFi* dapat dilihat pada tabel berikut:

Tabel 1. Spesifikasi *WiFi*.

Spesifikasi	Kecepatan	Frekuensi Band
802.11b	11 Mb/s	~ 2,4 GHz
802.11a	54 Mb/s	~ 2,4 GHz
802.11g	54 Mb/s	~2,4 GHz
802.11n	100 Mb/s	~ 5 GHz

Setiap perangkat *WiFi* atau biasa disebut *Access Point (AP)* yang terdapat dipasar sudah memiliki sertifikasi WiFi. Tiap-tiap perangkat ini memiliki standar yang berbeda-beda tergantung kemampuannya. Akan tetapi sebagian besar telah menggunakan standar 802.11 n.

Parameter-parameter *WiFi* secara umum adalah :

- Frekuensi adalah ukuran jumlah putaran ulang per peristiwa dalam selang waktu yang diberikan.
- Bandwidth adalah luas atau lebar cakupan frekuensi yang digunakan oleh sinyal dalam medium transmisi.
- Jangkauan Transmisi adalah untuk mengirimkan informasi dari suatu temoat ke tempat lain.

Komponen jaringan *WiFi* antara lain :

1. *Access Point*.

Access Point adalah suatu box tempat dimana kabel telepon dari pelanggan telepon terkoneksi. *Access Point* berfungsi sebagai pengatur lalu lintas data, sehingga memungkinkan banyak client dapat saling terhubung melalui jaringan (*Network*) sebagai *hub/switch* yang bertindak untuk menghubungkan jaringan lokal dengan jaringan wireless/nirkabel. Ada dua macam perangkat *Access Point*, yaitu *Dedicated Hardware Access Point*, adalah perangkat yang dibuat untuk bekerja sebagai *Access Point*, dan *Software Access Point* adalah komputer yang dilengkapi dengan wireless NIC dan software access point yang dibuat untuk menjadi access point. Ada 2 mode akses koneksi *WiFi* yaitu :

a. Ad-Hoc.

Mode koneksi ini adalah mode dimana beberapa komputer terhubung secara langsung, atau lebih dikenal dengan istilah peer-to-peer. Keuntungannya lebih murah dan praktis bila yang terkoneksi hanya 2 atau 3 komputer, tanpa harus membeli *Access Point*.

b. Infrastruktur.

Menggunakan *Access Point* yang berfungsi sebagai pengatur lalu lintas data, sehingga memungkinkan banyak client dapat saling terhubung melalui jaringan (*Network*).

2. Switch

System ini dapat digunakan untuk alat jaringan yang melakukan bridging transparan (penghubung segmentasi banyak jaringan dengan *forwarding* berdasarkan alamat MAC).

Ada beberapa jenis Switch, yaitu ATM Switch, ISDN Switch, DSLAM Switch, dan Ethernet Switch.

3. Web Server

Web Server merupakan software yang memberikan layanan data yang berfungsi menerima permintaan *HTTP* atau *HTTPS* dari klien yang dikenal dengan browser web dan mengirimkan kembali hasilnya dalam bentuk halaman-halaman web yang umumnya berbentuk dokumen HTML.

4. Radius

RADIUS (*Remote Authentication Dial-In User Service*) merupakan metode yang dianggap mudah diimplementasikan, sederhana, dan efisien. RADIUS adalah

sebuah *network* protokol keamanan komputer yang digunakan untuk membuat manajemen akses secara terkontrol pada sebuah jaringan besar, protokol membawa paket data, terdapat encapsulation di dalam paket data.

Karena perkembangannya sangat pesat dan kebutuhannya sangat dibutuhkan maka *WiFi* atau jaringan nirkabel banyak terdapat di banyak tempat, misalnya di kampus, tempat perbelanjaan, kafetaria, dan tempat area publik. Kekuatan jaringan *WiFi* tergantung kepada penerimaan sinyal oleh kartu jaringan wireless. Jika *WiFi* dipakai untuk hotspot untuk jaringan internet, sinyal yang menurun berakibatkan pada kecepatan browsing yang membuat masalah (Hary Nugroho, 2012).

2.2. Analisis Klaster.

Analisis Klaster mempunyai tujuan mengatur informasi dengan mengelompokkan beberapa objek berdasarkan kesamaan karakteristik tertentu. Teknik ini masuk dalam teknik *multivariate methods interdependent* yang mewakili hubungan interdependensi. Sehingga *clustering* memiliki tujuan mengelompokkan data dari beberapa pengamatan ke *subset* yang disebut klaster.

Tujuan analisis ini untuk menentukan struktur data dengan memasukkan observasi yang mirip ke satu kelompok berdasarkan korelasi antar objeknya. Hasilnya adalah dua observasi yang menunjukkan kesamaan. Langkah berikutnya adalah mengelompokkan objek berdasarkan karakteristik yang dimiliki sehingga didapatkan setiap objek yang paling dekat kesamaannya masuk ke dalam satu klaster. Dalam pengelompokan akan didapatkan homogenitas yang tinggi dalam satu kelompok dan heterogenitas yang tinggi antar kelompoknya. Manfaat analisis klaster adalah dapat memperringkas data.

Analisis cluster mengelompokkan sejumlah n objek berdasarkan p variabel yang secara relatif mempunyai kesamaan karakteristik di antara objek-objek tersebut, sehingga keragaman di dalam satu kelompok tersebut lebih kecil dibandingkan keragaman antar kelompok tersebut lebih kecil dibandingkan keragaman antar kelompok. Objek dapat berupa barang, jasa, tumbuhan, binatang dan orang (responden, konsumen, atau yang lainnya). Objek tersebut akan diklasifikasi ke

dalam satu atau lebih cluster (kelompok) sehingga objek-objek yang berada dalam satu cluster (kelompok) akan mempunyai kemiripan atau kesamaan karakter.

Analisis cluster merupakan suatu teknik yang dipergunakan untuk mengklasifikasi objek atau kasus dalam kelompok yang relatif homogen, yang disebut cluster. Objek dalam setiap kelompok cenderung mirip satu sama lain dan berbeda jauh (tidak sama) dengan objek dari cluster lainnya. Pengelompokan dilakukan berdasarkan kemiripan (*Similarity*) antar objek. Kemiripan diperoleh dengan meminimalkan jarak antar objek dalam kelompok (*WithinCluster*) dan memaksimalkan jarak antar kelompok (*Between-Cluster*) (Huberty, 2005).

2.3. K-Means.

K-Means merupakan salah satu algoritma *clustering* (X. Wu dan V. Kumar, 2009) Tujuan algoritma ini yaitu untuk membagi data menjadi beberapa kelompok. Algoritma ini menerima masukan berupa data tanpa label kelas. Hal ini berbeda dengan *supervised learning* yang menerima masukan berupa vektor $(-x-1, y1), (-x-2, y2), \dots, (-x-i, yi)$, di mana x_i merupakan data dari suatu data pelatihan dan y_i merupakan label kelas untuk x_i (S. Russel dan P. Norvig, 2010).

Pada algoritma pembelajaran ini, komputer mengelompokkan sendiri data-data yang menjadi masukannya tanpa mengetahui terlebih dulu target kelasnya (X. Wu dan V. Kumar, 2009). Pembelajaran ini termasuk dalam *unsupervised learning*. Masukan yang diterima adalah data atau objek dan k buah kelompok (cluster) yang diinginkan. Algoritma ini akan mengelompokkan data atau objek ke dalam k buah kelompok tersebut. Pada setiap cluster terdapat titik *centroid* yang merepresentasikan cluster tersebut.

Algoritma untuk melakukan *K-Means clustering* adalah sebagai berikut (P. N. Tan, M. Steinbach dan V. Kumar, 2005) :

1. Pilih K buah titik *centroid* secara acak
2. Kelompokkan data sehingga terbentuk K buah cluster dengan titik *centroid* dari setiap cluster merupakan titik *centroid* yang telah dipilih sebelumnya.

Proses pengelompokan data ke dalam suatu cluster dapat dilakukan dengan cara menghitung jarak terdekat dari suatu data ke sebuah titik *centroid*. Perhitungan jarak *Euclidean* dapat digunakan untuk menghitung jarak antar 2 buah data. Jarak

Euclidean adalah besarnya jarak suatu garis lurus yang menghubungkan antar objek. Misalkan dua objek yaitu A dengan koordinat (x_1, y_1) dan B dengan koordinat (x_2, y_2) maka jarak antar kedua objek tersebut dapat diukur dengan rumus $\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$. Ukuran jarak atau ketidaksamaan antar objek ke-i dengan objek ke-j, disimbolkan dengan d_{ij} dan $k=1, \dots, p$. Nilai d_{ij} diperoleh melalui perhitungan jarak kuadrat *Euclidean* sebagai berikut :

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (2.1)$$

Di mana:

D_{ij} = Jarak Kuadrat Euclidean antar objek ke-i dengan objek ke-j

P = Jumlah variabel cluster

X_{ik} = Nilai atau data dari objek ke-i pada variabel ke-k

X_{jk} = Nilai atau data dari objek ke-j pada variabel ke-k

(Everitt, 1993).

3. Perbaharui nilai titik *centroid*.

Untuk memperbarui titik *centroid* terdapat beberapa langkah :

a. Untuk setiap catatan, tentukan pusat kelompok terdekatnya dan tetapkan catatan tersebut sebagai kelompok anggota dari kelompok yang terdekat pusat kelompoknya.

b. Hitung BCV (*BetweenClusterVariation*) = jarak antar cluster

$$BCV = d(m_1, m_2) + d(m_2, m_3) + d(m_1, m_3) \quad (2.2)$$

c. Hitung WCV (*WithinClusterVariation*) = jarak antar anggota dalam cluster.

$$WCV = \sum_{i=1}^k \sum_{p=ci} d(p, m_i)^2 \quad (2.3)$$

d. Hitung rasio = $\frac{BCV}{WCV}$ (2.4)

e. Bandingkan rasio tersebut dengan rasio sebelumnya jika sudah ada. Jika rasio tersebut nilainya semakin besar, maka lanjutkan ke langkah ke-4. Namun jika tidak, hentikan prosesnya.

4. Ulangi langkah 2 dan 3 sampai nilai dari titik *centroid* tidak lagi berubah.

Pembaharuan suatu titik *centroid* dapat dilakukan dengan rumus berikut:

$$m_i = \frac{\sum c_i}{n} \quad (2.5)$$

Di mana:

m_i = titik *centroid* dari cluster ke-i

c = cluster

n = banyaknya data pada cluster

Ada beberapa kelebihan pada algoritma K-Means, yaitu (S. Russel dan P. Norvig, 2010) :

1. Mudah untuk diimplementasikan dan dijalankan.
2. Waktu yang dibutuhkan untuk menjalankan pembelajaran ini relatif cepat.
3. Mudah untuk diadaptasi.
4. Umum digunakan.

2.4. Purity

Menurut Manning, Raghavan, dan Schutz, purity merupakan salah satu ukuran untuk mengukur kualitas klustering berbasis *external criterion*. *External criterion* adalah metode untuk mengevaluasi seberapa baik hasil klustering dengan menggunakan sekumpulan kelas acuan sebagai wakil penilaian pengguna, dimana kelas acuan ini diperoleh dari hasil penilaian manusia. Evaluasi ini tidak menggunakan label kelas pada kelas acuan, tetapi hanya menggunakan hasil partisi pada kelas acuan. Untuk menghitung purity dapat menggunakan rumus berikut:

$$\text{Purity}(\Omega, C) = \frac{1}{N} \sum_k \frac{\max_j |\omega_k \cap C_j|}{j} \quad (2.13)$$

Dimana :

Ω = $\{w_1, w_2, \dots, w_k\}$ adalah kumpulan cluster.

C = $\{C_1, C_2, \dots, C_j\}$ adalah kumpulan kelas acuan.

W_k = adalah kumpulan objek pada cluster w_k .

C_j = adalah kumpulan objek pada kumpulan kelas acuan C_j

J = jumlah objek pada kumpulan kelas

N = Jumlah klaster

Jika nilai purity semakin mendekati 1 menunjukkan bahwa klustering memberikan hasil yang semakin baik. Sedangkan jika nilai purity semakin mendekati 0, maka klustering memberikan hasil yang semakin buruk (Manning, 2008).

BAB 3. TUJUAN DAN MANFAAT

3.1 Tujuan.

Tujuan penelitian ini adalah menghasilkan kelompok/cluster sebagai rekomendasi pembatasan kuota bandwidth yang diberikan ke tiap user id dalam satu semester dengan menggunakan metode pengelompokkan K-Means, untuk optimalisasi penggunaan internet

3.2 Manfaat dan Kontribusi terhadap ilmu Pengetahuan

Manfaat dan kontribusi terhadap ilmu pengetahuan yang diperoleh dari hasil penelitian ini adalah:

1. Publikasi hasil penelitian tentang penggunaan metode *K-Means* untuk optimalisasi penggunaan internet dengan pembagian kuota *bandwidth* berdasarkan kelompok pengguna
2. Dari hasil penelitian ini yang berupa data klaster, hasil penelitian dapat memperbanyak manfaat pengelompokkan dengan metode K-Means yang dapat digunakan juga untuk pengelolaan jaringan.

BAB 4. METODE PENELITIAN

Dalam penelitian ini metodologi dapat dijelaskan sebagai berikut :

1. Studi pendahuluan.
2. Observasi/studi pustaka.
3. Perumusan masalah.
4. Pembuatan Sistem.
5. Pengujian.
6. Implementasi.

4.1. Studi Pendahuluan.

Pada bagian ini dilakukan pemetaan terhadap obyek dan hal-hal yang berkaitan obyek penelitian. Pengumpulan informasi tentang variable penelitian dan cara pengambilan data tiap variabel yang akan dikelompokkan dengan metode kluster K-Means.

4.2. Observasi/Studi Pustaka.

Observasi dilakukan di Kampus Universitas Muhammadiyah magelang, untuk melihat situasi dan kondisi yang ada di lokasi. Pengumpulan informasi di lapangan dilanjutkan dengan informasi dari buku, tulisan ilmiah, internet, dan media ilmiah lainnya.

4.3. Perumusan Masalah.

Tahap selanjutnya adalah menentukan rumusan masalah utama dari sekian banyak masalah hasil observasi. Penentuan masalah didukung dengan hasil studi pustaka tentang bagaimana kemungkinan untuk mengatasi masalah tersebut.

4.4. Pembuatan Sistem.

Langkah-langkah yang akan dilakukan dalam pembuatan sistem adalah sebagai berikut:

3.4.1. Analisis Masalah.

- a. Universitas Muhammadiyah Magelang mempunyai jumlah mahasiswa aktif pada semester ganjil 2015/2016 sebanyak 4777 mahasiswa dari 16 program

studi. Untuk mendukung suasana akademik maka mahasiswa dapat menggunakan fasilitas internet yang dapat diakses selama 24 jam.

- b. Perbedaan kebutuhan dalam penggunaan internet memberikan pengaruh dalam mengakses internet. Sehingga perlu adanya batasan kuota yang diberikan. Untuk membatasi kuota tersebut maka diperlukan suatu pengelompokan penggunaan internet, dimana akan ditentukan menjadi 3 kelompok yaitu kelompok pengguna internet besar, kelompok pengguna internet sedang, dan pengguna internet kecil.
- c. Dari hasil kelompok klaster tersebut, diharapkan dapat menjadi rekomendasi untuk berapa besar kuota yang diberikan ke tiap-tiap user id. Karena kebutuhan internet dari tiap-tiap pengguna berbeda.

3.4.2. Analisis Data Sistem.

Data yang digunakan dalam penelitian ini yang akan digunakan untuk pengelompokan adalah data total upload, data total download, dan data berapa banyak tiap-tiap pengguna tersebut login. Karena kadang ada user id yang banyak melakukan upload, dan ada user id yang banyak melakukan download saja, dan banyak login yang dilakukan mengindikasikan bahwa pengguna banyak menggunakan internet.

3.4.3. Analisis Kebutuhan Fungsional.

Kebutuhan fungsional adalah jenis kebutuhan yang berisi proses-proses apa saja yang nantinya dilakukan oleh sistem. Kebutuhan fungsional juga berisi informasi-informasi apa saja yang harus ada dan dihasilkan oleh sistem. Berikut kebutuhan fungsional yang terpadat pada sistem yang dibangun :

1. Mengimplementasikan K-Means *clustering* pada data upload, download, dan banyak login pengguna internet di Universitas Muhammadiyah Magelang dalam rentang waktu satu semester.
2. Algoritma yang digunakan dalam melakukan *clustering* ini adalah algoritma K-Means.
3. Objek yang dikelompokkan adalah data upload, data download, data banyak login dari tiap pengguna internet.
4. Hasil yang didapatkan dari implementasi ini adalah kelompok pengguna internet kedalam pengguna internet besar, sedang, dan kecil. Hasil

pengelompokan tersebut dapat dijadikan rekomendasi untuk pembatasan kuota internet semester berikutnya.

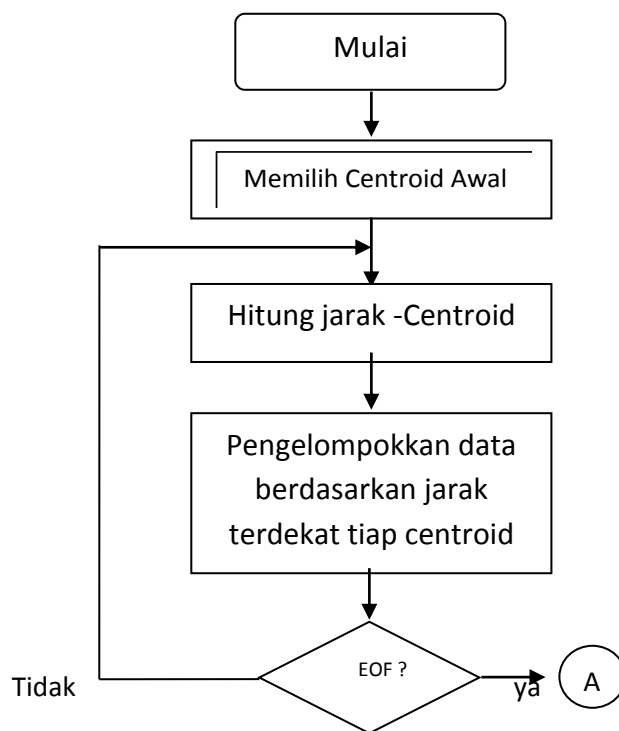
3.4.4. Analisis Kebutuhan Non Fungsional.

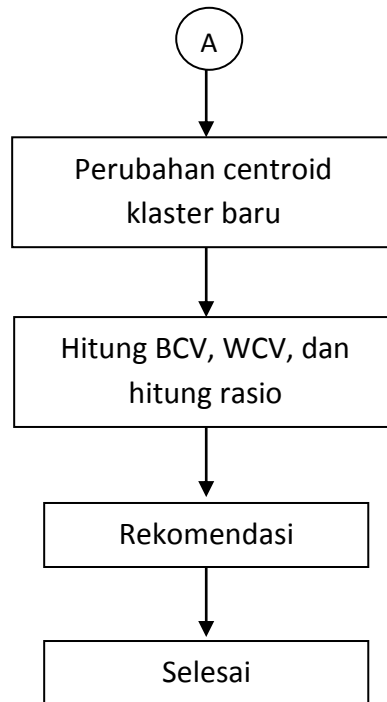
Kebutuhan ini adalah tipe kebutuhan yang berisi properti yang dibutuhkan untuk mendukung proses penelitian. Berikut adalah kebutuhan non fungsional yang dibutuhkan :

1. Digunakan pada sistem operasi *Microsoft Windows*, minimal *Microsoft Windows 7*.
2. Implementasi algoritma dilakukan dengan menggunakan *Matlab&SimulinkMathworks 2013*, digunakan untuk implementasi algoritma *K-Meansclustering*.
3. Perangkat android Wifi digunakan untuk monitoring trafik internet.
4. Spesifikasi komputer yang digunakan *Processor Intel Core 2 Duo 2,2 GHz*, Memori 4 GB, Kartu Grafik *NVidiaGeForce G105M Cuda512 MB*.
5. Hardisk eksternal digunakan untuk backup data.

3.4.5. Perancangan Sistem.

Flowchart penelitian dapat dilihat pada gambar berikut:





4.5. Pengujian.

Rencana pengujian dilakukan untuk menguji kualitas kluster. Pengujian akan dilakukan menggunakan metode *Purity*. Sistem akan menghasilkan tiga kluster, dimana tiap kluster terdapat beberapa anggota. Dengan melihat besar nilai puritinya, jika nilainya mendekati 1 maka kualitas klasternya semakin baik. Sehingga akan didapatkan hasil pengelompokkan dimana hasil tiap kelompok akan mempunyai data yang mendekati masing-masing pusat kluster yang terbentuk.

4.6. Implementasi

Tahap akhir dari penelitian ini adalah mengimplementasikan sistem dengan cara mengkluster data upload, download, dan banyak login sehingga menghasilkan kelompok anggota pengguna besar, sedang, dan kecil dimana hasil tersebut dapat menjadi rekomendasi untuk kebijakan pembatasan kuota secara adil.

BAB 5 HASIL DAN PEMBAHASAN

5.1 Hasil Pengelompokkan Pengguna Internet

Data pengguna dalam penelitian terdiri dari 3 macam, yaitu data login, data download dan data upload dalam satuan giga bit untuk setiap pengguna. Data tersebut diperoleh dari hasil monitoring penggunaan internet untuk periode waktu satu sampai tiga bulan. Sumber data pengguna tersebut diperoleh dari sistem pengelolaan dan monitoring user pengguna internet Universitas Muhammadiyah Magelang.

Pengumpulan data dalam periode satu sampai tiga bulan tercatat pada system sebanyak 16.433 user. Kemudian setelah dilakukan proses normalisasi, diperoleh 3000 data pengguna aktif. Setiap data pengguna ini terdiri dari data login, data upload dan data download. Hasil pengelompokkan dengan menggunakan metode K Means adalah seperti terlihat pada Tabel 1.

Tabel 1. Hasil Kelompok dan Nilai Purity

Jumlah Data	Kelompok 1	Kelompok 2	Kelompok 3	Purity
3000	2328	104	568	0,85

Jumlah total data hasil klasifikasi diatas tentunya menjadi beban komputasi tersendiri pada saat proses *clustering* berlangsung. Dengan beban komputasi yang sangat besar tersebut, proses clustering akan berjalan sangat lama dan dapat mengakibatkan proses tidak pernah konvergen. Oleh sebab itu, masih dalam tahap klasifikasi data, banyaknya data yang digunakan diambil sesuai dengan proporsi dari jumlah keseluruhan pengguna internet setiap kelompok penggunaan.

Klasifikasi dan pemisahan data untuk data pengujian memegang peranan penting dalam membentuk karakteristik dari *k-means clustering* yang dihasilkan. Oleh sebab itu, dengan diperolehnya data pengelompokkan dan hasil pengelompokkan tersebut dilakukan pengujian maka tahapan penting pertama dari proses penelitian ini sudah dapat diselesaikan. Selanjutnya dengan data pengelompokkan tersebut akan dijadikan rekomendasi berdasarkan hasil kelompoknya.

Hasil 5 data dari tiap klaster dapat dilihat di tabel 2 berikut:

Tabel 2. Hasil 5 data tiap klaster

Kelompok	Urutan Data	Upload	Download	Jumlah Login
Kelompok 1	Data ke 1	0.0014	0.0192	7
	Data ke 3	0.0348	0.5316	23
	Data ke 4	0.2582	2.3428	78
	Data ke 5	0.1387	2.6431	57
	Data ke 6	0.1608	2.0926	96
	Kelompok 2	Data ke 2	1.3429	19.9585
Data ke 43		1.206	11.4907	668
Data ke 50		1.7633	43.8502	667
Data ke 99		1.0182	12.435	650
Data ke 128		0.8426	10.8729	766
Kelompok 3		Data ke 11	0.2457	7.1819
	Data ke 15	1.8786	11.8398	304
	Data ke 16	2.9313	62.1799	482
	Data ke 18	0.2666	2.7826	221
	Data ke 28	1.5569	44.3769	347

Dari 5 hasil dari tiap kelompok diatas, dapat dilihat bahwa kelompok trafik internet terkecil adalah kelompok 1, kemudian trafik internet sedang adalah kelompok 3, dan yang terbesar adalah kelompok 2. Sehingga dari hasil kelompok ini rekomendasi pembatasannya secara urut adalah kelompok 2 (104 anggota), Kelompok 3 (568 anggota) dan kelompok 1 (2328 anggota). Besar pembatasan yang akan dilakukan merupakan kebijakan dari Biro PDSI dengan melihat besar bandwidth yang digunakan.

5.2 Hasil Pengujian Kualitas Pengelompokkan dengan Variasi Data

Dari data yang digunakan di atas, yaitu 3000 data dapat digunakan untuk mengetahui perilaku hasil kelompok/klaster jika data yang digunakan divariasikan. Variasi data yang akan digunakan adalah variasi 50, 100, 150, 200, 250, 300, dan 400 data. Dari tiap banyak data tersebut akan divariasikan sebanyak 5 variasi.

Sehingga hasil dari tiap kelompok merupakan rata-rata dari tiap 5 variasi data tersebut. Hasil rata-rata kelompok dari tiap variasi dapat dilihat di Tabel 3.

Tabel 3. Rata-Rata Hasil Klaster Tiap Variasi Data.

Variasi Data	3 Kelompok			4 Kelompok				5 Kelompok				
	Kel 1	Kel 2	Kel 3	Kel 1	Kel 2	Kel 3	Kel 4	Kel 1	Kel 2	Kel 3	Kel 4	Kel 5
50 data	32.6	5.6	11.8	28.6	4.2	10.8	6.4	23.4	6.6	12.2	5.2	2.6
100 data	67.4	8.6	24	55.4	12.2	23.6	8.8	44.2	17.4	24.2	10	4.2
150 data	105	10.2	34.6	84.8	21	37.6	6.6	64.4	25.8	35.2	18.2	6.4
200 data	147	11	42.2	114	28	51	6.8	95.4	30.4	50.8	16.8	6.6
250 data	188	12.4	49.6	146	34	61.2	8.6	123	36.8	60.4	22.4	7.2
300 data	229	14.2	57	171	40.2	77.6	11.4	148	42.8	72.4	26.4	10
350 data	0.77	0.95	0.81	208	45.2	83.8	13.4	179	45.6	89.8	25.6	10.2
400 data	0.77	0.96	0.81	255	41.6	90.8	12.6	205	55	99.8	29.2	11.4
3000 data	2328	104	568	2238	163	575	24	1752	125	698	401	24

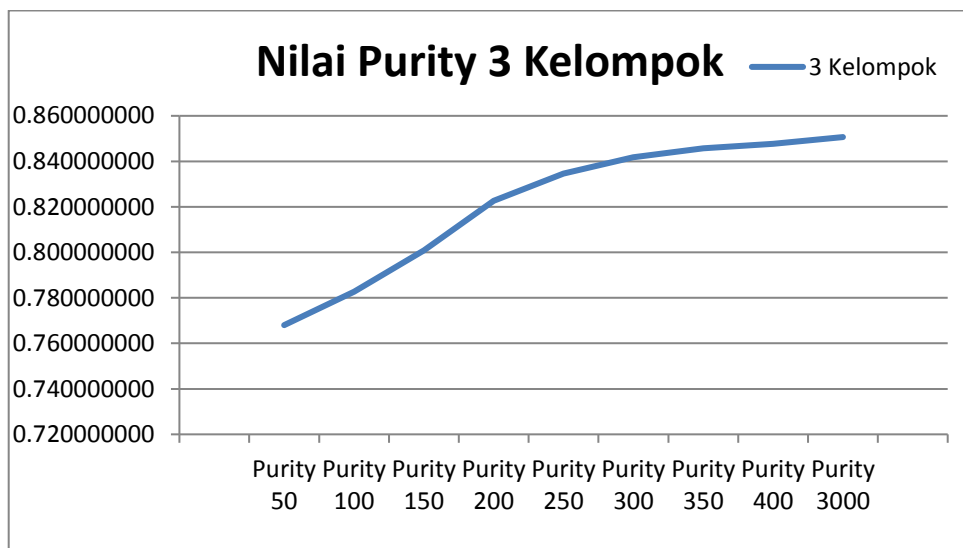
Dari tabel diatas merupakan rata-rata dari 5 kelompok data hasil klaster dari tiap-tiap variasi datanya. Hasil dari klaster terlihat divariasikan menjadi tiga jenis kelompok, yaitu hasil klaster 3 kelompok, 4 kelompok, dan 5 kelompok. Dari hasil kelompok tersebut diharapkan dapat mengetahui kualitas klasternya bagaimana. Untuk mengetahui kuitas hasil klasternya digunakan metode Purity. Kualitas klaster akan lebih baik jika hasil puritynya mendekati angka satu. Berikut hasil purity dari hasil klaster pada tabel 3 di atas dapat dilihat pada tabel 4. Tabel nilai purity.

Tabel 4. Nilai Purity

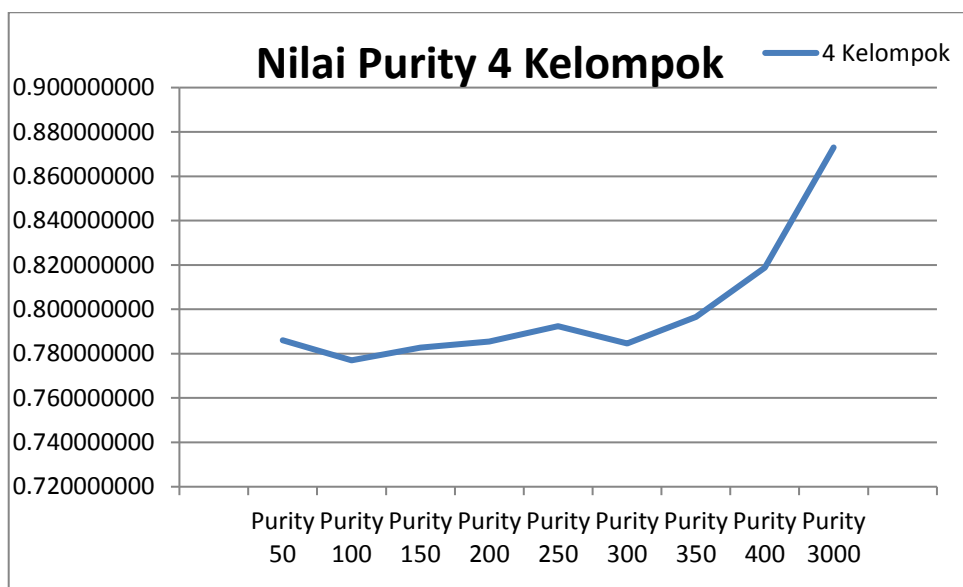
Variasi Data	3 kelompok	4 kelompok	5 kelompok
50 data	0.768000000	0.786000000	0.800000000
100 data	0.782666667	0.777000000	0.800000000
150 data	0.800888889	0.782666667	0.800000000
200 data	0.822666667	0.785500000	0.800000000

250 data	0.834666667	0.792400000	0.800000000
300 data	0.841777778	0.784666667	0.800000000
350 data	0.845714286	0.796571429	0.804342857
400 data	0.847666667	0.818750000	0.804600000
3000 data	0.850666667	0.873000000	0.833600000

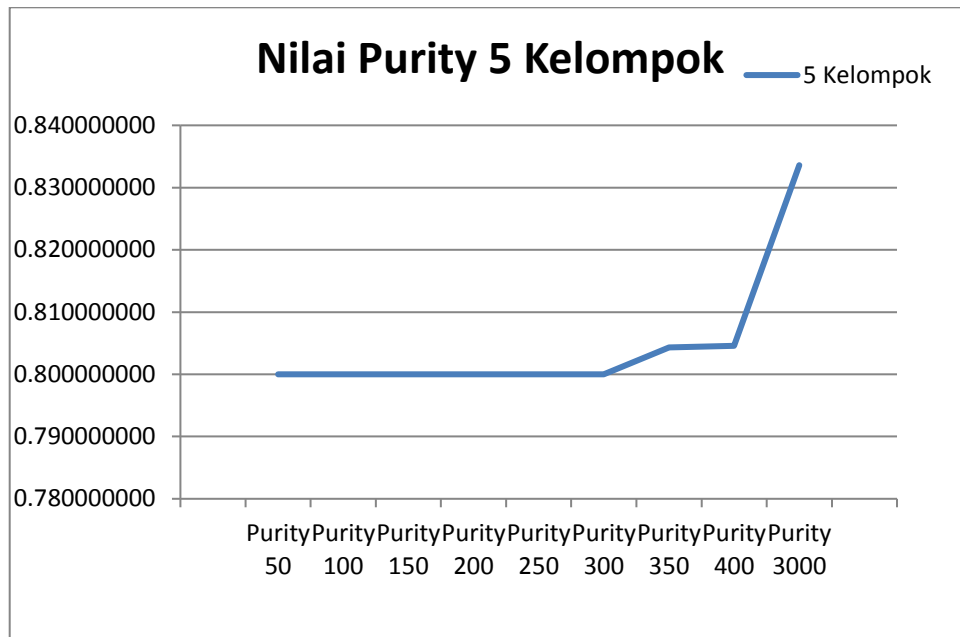
Dari tabel 4. Nilai Purity, maka jika ditampilkan dalam bentuk grafik dapat dilihat pada Gambar 2, Gambar 3, dan Gambar 4, yaitu:



Gambar 3. Grafik Nilai Purity kelompok 3



Gambar 4. Grafik Nilai Purity kelompok 4



Gambar 5. Grafik Nilai Purity kelompok 5

Dari grafik di atas terlihat bahwa terjadi perubahan secara linier hubungan antara banyak data bertambah ikut merubah semakin mendekati satu untuk nilai puritinya.

Besar purity yang dihasilkan untuk 3 cluster adalah 0,85. Nilai tersebut jika diubah dalam persen adalah 85%. Nilai tersebut lebih dari 0,5 dan mendekati nilai 1, sehingga dapat dikatakan kualitas kelompok yang dihasilkan dengan metode K Means kualitas kelompoknya bagus. Sehingga hasil anggota tiap kelompok mempunyai kemiripan yang besar (homogen) dan dari antar kelompok mempunyai perbedaan yang besar (heterogen).

BAB 6 KESIMPULAN

Kesimpulan yang dapat diperoleh dari hasil penelitian sampai laporan perkembangan ini disusun adalah:

1. Tahap klasifikasi dan pemisahan data untuk mendapatkan data pengujian, memegang peranan penting dalam membentuk karakteristik dan kemampuan dari *neural network* yang dihasilkan.
2. Untuk menghindari beban komputasi yang terlalu besar pada saat proses pengujian *k-means clustering* berlangsung yang berdampak pada. Pada penelitian ini dari 9000 data yang terdiri masing-masing 3000 data login, 3000 data download dan 3000 data upload, akan dikelompokkan lagi sesuai fungsinya untuk mengurangi beban komputasi. Karena beban komputasi yang terlalu besar akan berakibat proses clustering data akan berjalan sangat lama dan dapat mengakibatkan proses tidak pernah konvergen.
3. Hasil kelompok yang didapatkan dengan metode K Means adalah kelompok 1 (2328 anggota), kelompok 2 (104 anggota), dan kelompok 3 (568 anggota).
4. Besar nilai purity 0,85 sehingga hasil kelompok dalam satu kelompok datanya homogen dan hasil antar kelompok datanya heterogen.
5. Urutan rekomendasi pembatasan bandwidth nya berdasarkan urutan trafik internet terbesar yaitu kelompok 2 trafik internet terbesar, kelompok 3 trafik internet sedang, dan kelompok 1 trafik internet terkecil.

DAFTAR PUSTAKA

- Huberty, C. J., Jordan, E. M., & Brandt, W. C. (2005). Cluster Analysis In Higher Education Research. In J. C. Smart (Ed), Higher Education: Handbook of Theory and Research (Vol. 20, pp. 437-457). Great Britain: Springer.
- Manning, C. D., Raghavan, P., Schutze, H. (2008). Introduction to Information Retrieval. Cambridge University Press. New York.
- Nugroho, Hary., Siagian Sartika, A. Analisis Bandwidth Jaringan WiFi Studi Kasus Telkom Jakarta Pusat. Jurnal ICT Penelitian dan Penerapan Teknologi. Akademi Telkom Sandhy Putra Jakarta, 2012. Jakarta.
- Russel, Stuart., Norvig, Peter. Artificial Intelligence A Modern Approach Third Edition. Pearson Education, Inc. 2010. New Jersey.

Lampiran 1: Susunan organisasi tim peneliti dan pembagian tugas

No	Nama / NIDN	Instansi Asal	Bidang Ilmu	Alokasi Waktu (jam/minggu)	Uraian Tugas
1	Mukhtar Hanafi, ST., MCs. / 0602047502	universitas Muhammadiyah Magelang	Teknik Informatika	10	Menganalisa dan memonitor trafik internet
2	Auliya Burhanuddin, S.Si / 0630058202	universitas Muhammadiyah Magelang	Ilmu Komputer	8	Menganalisa dan merancang klastering dengan K-Means.

1. Biodata Ketua Tim Peneliti

A. Identitas Diri

1	Nama Lengkap (dengan gelar)	Mukhtar Hanafi, ST., MCs.
2	Jenis Kelamin	L
3	Jabatan Fungsional	Asisten Ahli
4	NIP/NIK/Identitas Lainnya	057508191
5	NIDN	0602047502
6	Tempat dan Tanggal Lahir	Magelang, 2 April 1975
7	E-mail	hanafi@ummgl.ac.id
9	Nomor Telepon / HP	0817466231
10	Alamat Kantor	Kampus 2 : Jl. Mayjend Bambang Soegeng Mertoyudan Magelang Jawa Tengah 56172
11	Nomor Telepon/Faks	(0293)362082 / (0293) 361004
12	Lulusan yang telah dihasilkan	S-1 = 54 orang; S-2 = - orang; S-3 = - orang.
13	Mata Kuliah yang Diampu	<ol style="list-style-type: none"> 1. Pemrograman 1 2. Pemrograman 2 3. Pemrograman 3 4. Pemrograman Berorientasi Obyek 5. Artificial Intelligence (AI)

B. Riwayat Pendidikan

	S – 1	S – 2	S – 3
Nama Perguruan Tinggi	Univ. Muhammadiyah Yogyakarta	Universitas Gadjah Mada	
Bidang Ilmu	Teknik Elektro	Ilmu Komputer	
Tahun Masuk	1994	2007	
Judul Skripsi/Thesis/Desertasi	Penerapan Pengendali Logika Fuzzy pada Pengendali PID dengan Waktu Tunda	Aplikasi Sitem Hybrid Fuzzy Neural Network (FNN) pada Sistem Kontrol Suspensi Aktif	
Nama Pembimbing/Promotor	1. Adhi Susanto, M.Sc., Ph.D. 2. Ir. Fathul Qodir	1. Drs. Retantyo Wardoyo, M.Sc., Ph.D.	

C. Pengalaman Penelitian Dalam 5 Tahun Terakhir

(Bukan Skripsi, tesis, maupun Disertasi).

No.	Tahun	Judul Penelitian	Pendanaan	
			Sumber*	Jml (Juta Rp)
1	2013	<i>Artificial Neural Network</i> Berbasis Indeks Prestasi Akademik Guna Prediksi Keberhasilan Studi Mahasiswa	Dikti	14,5

D. Pemakalah Seminar Ilmiah (*Oral Presentation*) dalam 5 Tahun Terakhir

No	Nama Pertemuan Ilmiah / Seminar	Judul Artikel Ilmiah	Waktu dan Tempat
1	Seminar Nasional Informatika (SEMNASIF) 2012	<i>Aplikasi Fuzzy Neural Network (FNN) pada Sistem Kontrol dengan Waktu Tunda</i>	30 Juni 2012, Univ. Pembangunan Nasional (UPN) 'Veteran' Yogyakarta
2	Seminar Nasional Informatika (SEMNASIF) 2013	Meningkatkan Kemampuan Pengenalan Pola Sinyal Dengan Optimalkan <i>Rules</i> Pada <i>Fuzzy Neural Network</i>	18 mei 2013, Univ. Pembangunan Nasional (UPN) 'Veteran' Yogyakarta
3	Konferensi Nasional System Informasi (KNSI) 2014	Prediksi Keberhasilan Studi Mahasiswa Menggunakan <i>Artificial Neural Network</i> Berbasis Indeks Prestasi Akademik	27 Februari 2014, STIMIK Dipanegara Makasar

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggungjawabkan secara hukum. Apabila di kemudian hari ternyata dijumpai ketidaksesuaian dengan kenyataan, saya sanggup menerima sanksi.

Demikian biodata ini saya buat dengan sebenarnya untuk memenuhi salah satu persyaratan dalam pengajuan Hibah Penelitian Dosen Pemula

Magelang, 2 Mei 2016

pengusul,



Mukhtar Hanafi, ST., M.Cs

2. Biodata Anggota Tim Peneliti

A. Identitas Diri

1	Nama Lengkap (dengan gelar)	Auliya Burhanuddin, S.Si
2	Jenis Kelamin	L
3	Jabatan Fungsional	-
4	NIP/NIK/Identitas Lainnya	098106037
5	NIDN	0630058202
6	Tempat dan Tanggal Lahir	Surakarta, 30 Mei 1982
7	E-mail	Burhan_a23@ummgl.ac.id
9	Nomor Telepon / HP	085743469726
10	Alamat Kantor	Kampus 2 : Jl. Mayjend Bambang Soegeng Mertoyudan Magelang Jawa Tengah 56172
11	Nomor Telepon/Faks	(0293)362082 / (0293) 361004
12	Lulusan yang telah dihasilkan	S-1 = 54 orang; S-2 = - orang; S-3 = - orang.
13	Mata Kuliah yang Diampu	6. Logika Informatika. 7. Teori Bahasa dan Otomata 8. Fisika Dasar 9. Aljabar Linier 10. Algoritma dan Pemrograman

B. Riwayat Pendidikan

	S – 1	S – 2	S – 3
Nama Perguruan Tinggi	Universitas Sebelas Maret	Universitas Gadjah Mada	
Bidang Ilmu	Mipa Fisika	Kecerdasan Buatan	
Tahun Masuk	2001	2007	
Judul Skripsi/Thesis/Desertasi	Kolektor Surya Plat datar dengan variasi jarak kaca dan sudut kemiringan	Perbandingan Metode Pengelompokan <i>Fuzzy C-Means</i> (FCM) dengan <i>Fuzzy Possibilistic C-Means Standar</i> (FPCM) pada komoditi Industri dan Kerajinan Berdasarkan Volume dan	

		Nilai Ekspor (Studi Kasus di Kabupaten Sleman).	
Nama Pembimbing/Promotor	3. Prof. Ir. Ari Handono Ramelan, M.Sc(hons)., Ph.D 4. Drs. Harjana, M.Sc., Ph.D	3. Drs. Retantyo Wardoyo, M.Sc., Ph.D.	

C. Pengalaman Penelitian Dalam 5 Tahun Terakhir

(Bukan Skripsi, tesis, maupun Disertasi).

No.	Tahun	Judul Penelitian	Pendanaan	
			Sumber*	Jml (Juta Rp)
1	2012	Helm Las Listrik Otomatis untuk Peningkatan Keselamatan Kerja pada Proses Pengelasan	LP3M UMM	4,76
2	2013	Aplikasi Jaringan Syaraf Tiruan untuk Prediksi Keberhasilan Studi	Penelitian Dosen Muda	14,5

D. Pemakalah Seminar Ilmiah (*Oral Presentation*) dalam 5 Tahun Terakhir

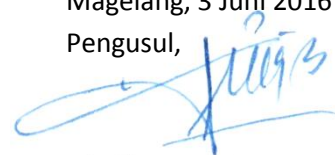
No	Nama Pertemuan Ilmiah / Seminar	Judul Artikel Ilmiah	Waktu dan Tempat
1	Proceedings Konferensi Nasional Sistem Informasi (KNSI)	Implementasi Cloud Computing pada VPN Berbasis Komunitas Sekolah	Stimik Bumigora Mataram, 14-15 Pebruari 2013
2	Proceedings Konferensi Nasional Sistem Informasi (KNSI)	Penerapan Algoritma Ant Colony System untuk Penyelesaian Travelling System Problem.	Universitas Klabat Airmadidi, minahasa utara Sulawesi utara, 30 januari 2015
3	Konferensi Nasional System Informasi (KNSI) 2014	Prediksi Keberhasilan Studi Mahasiswa Menggunakan <i>Artificial Neural Network</i> Berbasis Indeks Prestasi Akademik	27 Februari 2014, STIMIK Dipanegara Makasar

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggungjawabkan secara hukum. Apabila di kemudian hari ternyata dijumpai ketidaksesuaian dengan kenyataan, saya sanggup menerima sanksi.

Demikian biodata ini saya buat dengan sebenarnya untuk memenuhi salah satu persyaratan dalam pengajuan Hibah Penelitian Dosen Pemula

Magelang, 3 Juni 2016

Pengusul,



Auliya Burhanuddin, S.Si

