

## Clustering Kegiatan Pengguna pada Media Sosial dengan Algoritma *Simple K-Means*

Rivki Maulana<sup>1</sup>, Iqbal Gibran<sup>2</sup>, Desti Fitriati<sup>3</sup>

Fakultas Teknik Program Studi Teknik Informatika  
Universitas Pancasila

email: [maulanarivki8@gmail.com](mailto:maulanarivki8@gmail.com)<sup>1</sup>, [iqbalgibraaan@gmail.com](mailto:iqbalgibraaan@gmail.com)<sup>2</sup>,  
[desti.fitriati@univpancasila.ac.id](mailto:desti.fitriati@univpancasila.ac.id)<sup>3</sup>

Jl. Srengseng Sawah, Jagakarsa, Jakarta Selatan, DKI Jakarta, 12640, Indonesia

### Abstrak

Media sosial merupakan salah satu jenis media komunikasi yang populer. Terdapat berbagai macam kegiatan yang bisa dilakukan, mulai dari mengirim pesan, berbelanja dan lainnya. Sehingga mampu memudahkan seseorang untuk berkomunikasi dengan orang lain. Dari berbagai macam kegiatan yang biasa dilakukan di sosial media beberapa pengguna cenderung melakukan kegiatan yang berubah, pada penelitian ini dipilih beberapa jenis kegiatan yang umum dilakukan pada media sosial. Penelitian ini bertujuan untuk mengelompokkan kegiatan yang memiliki sifat yang sama dan biasa digunakan pengguna media sosial pada lingkungan Mahasiswa. Dari data hasil pengelompokan kegiatan tersebut nantinya bisa dilihat fitur apa saja yang bermanfaat bagi pengguna media sosial dan data tersebut bisa menjadi bahan acuan atau evaluasi para software development dalam mengembangkan aplikasi media sosial. Penelitian ini menggunakan metode *Clustering* untuk mengelompokkan data - data yang telah didapat, algoritma yang digunakan yaitu Algoritma *Simple K-Means*. Pengelompokan data pada penelitian ini dilakukan dengan menggunakan tools bernama WEKA. Pada penelitian ini, didapatkan 3 *cluster* yang berbeda sesuai dengan ciri khas masing masing responden. Dimana 50% responden termasuk dalam Pengguna Aktif, 27% responden termasuk Pengguna Sedang dan 23% responden termasuk Pengguna Normal.

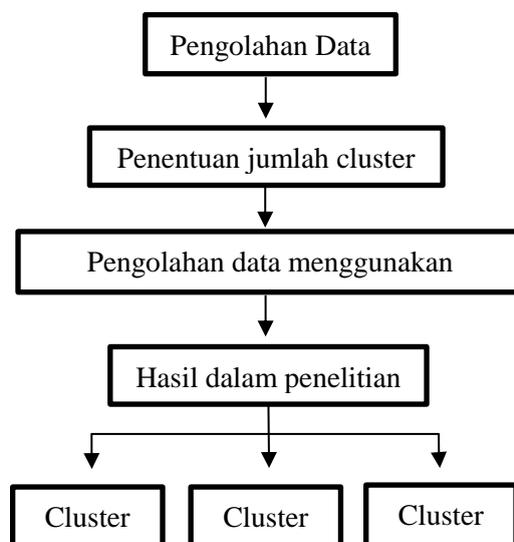
Kata kunci: Media Sosial, Algoritma *Simple K Means*, *clustering*, WEKA

### 1 PENDAHULUAN

Media komunikasi saat ini sudah banyak tersedia baik untuk komunikasi suara, pengiriman pesan, bahkan komunikasi dengan menggunakan teknologi video call. Tujuan tersedianya media komunikasi adalah sebagai alat untuk mempermudah penyampaian informasi dengan cepat, dan mudah. Salah satu bentuk media komunikasi adalah media sosial. Pada dasarnya media sosial dibuat untuk memudahkan orang berkomunikasi, berpartisipasi, saling berbagi dan membentuk jaringan secara online.

Dalam penelitian ini kami melakukan pengelompokan berdasarkan jenis kegiatan yang biasa dilakukan pengguna di media sosial pada kalangan mahasiswa. pengelompokan dibuat menjadi 3 kategori terdiri dari aktif, normal dan pasif, dengan adanya pengelompokan ini maka akan diperoleh peluang untuk mengetahui fitur apa yang digemari dan bisa menjadi bahan acuan atau evaluasi para software development dalam mengembangkan aplikasi media sosial berdasarkan dengan kelompok kegiatan yang dihasilkan

## 2 METODOLOGI PENELITIAN



Gambar 1. Bagan Alir

### 2.1 Pengolahan Data

Pada penelitian ini, data yang sudah diperoleh lalu diolah lagi agar bisa diproses hingga tahap akhir proses. Hasil dari pengolahan data adalah data *training set* akhir dengan proses data mining.

Berikut adalah tahapan – tahapan Pengolahan Data :

#### A. Data Primer

Data primer adalah data yang didapat peneliti dari sumber pertama baik individu atau perseorangan seperti hasil wawancara atau pengisian kuesioner yang biasa dilakukan oleh peneliti. Pada penelitian ini data primer berupa data yang berkaitan dengan kegiatan yang biasa dilakukan di media sosial. Data tersebut kami dapat dari hasil survey kuesioner dilingkungan mahasiswa.

#### B. Pembersihan Data

Pembersihan Data adalah teknik yang biasa dipakai untuk membuang data yang tidak sesuai dengan penelitian. Pada penelitian ini kami membuang beberapa data yang telah di dapat dari hasil survei, hal ini dilakukan untuk meningkatkan hasil penelitian

#### C. Seleksi Data

Selepas data telah terkumpul semua, maka dilakukan proses seleksi. Tahapan ini digunakan untuk memilih data mana saja yang akan diambil dan dijadikan masukan pada proses mining. Hal ini dilakukan agar data menjadi lebih mudah diolah dan meningkatkan keakuratan.

#### D. Transformasi Data

Pada tahapan ini hanya mengubah data yang telah terpilih dari hasil survey sebelumnya menjadi format yang sesuai dengan Aplikasi WEKA agar data dapat diolah.

### 2.2 Clustering

Salah satu teknik yang dikenal dalam *Data Mining* yaitu *clustering*. Pengertian *clustering* keilmuan dalam *Data Mining* adalah pengelompokan sejumlah data atau objek ke dalam *cluster*

(*group*) sehingga setiap dalam *cluster* tersebut akan berisi data yang semirip mungkin dan berbeda dengan objek dalam *cluster* yang lainnya. Analisis *Cluster* merupakan teknik multivariat yang mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya. Analisis *Cluster* mengklasifikasi objek sehingga setiap objek yang paling dekat kesamaannya dengan objek lain berada dalam *cluster* yang sama.

### 2.3 Algoritma *Simple K-means*

Simple K-Means adalah suatu metode penganalisaan data atau metode data mining yang melakukan proses pemodelan tanpa supervisi (*unsupervised*) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. Metode ini berusaha untuk meminimalkan variasi antar data yang ada didalam suatu cluster dan memaksimalkan variasi dengan data yang ada di *cluster* lainnya.

### 2.4 WEKA

WEKA adalah aplikasi *machine learning* praktis yang biasa digunakan untuk penggalian informasi dari Data Mining menggunakan algoritma seperti *Klasifikasi*, *Regresi*, dan *Clustering*.

Pada penelitian ini, ditentukan jumlah *cluster* yaitu sebesar 3 *cluster*. Hal ini dilakukan agar didapatkan hasil maksimal dalam pengelompokan data. Lalu data diolah dengan menggunakan bantuan tools bernama WEKA.

## 3 HASIL DAN PEMBAHASAN

### 3.1 Data

Pada tahap pertama data yang diperoleh dari hasil survei lalu diberikan atribut variable kegiatan berdasarkan abjad agar data lebih mudah diolah.

Tabel 1: Atribut data

Kategori	Keterangan Kegiatan Pengguna
A	Mengirim Pesan
B	Menyukai Kiriman
C	Komentar pada postingan
D	Update status
E	Bermain game
F	Membaca berita
G	Berbelanja
H	Berjualan

Pada tahap kedua data diubah dalam bentuk angka 1 dan 0, dimana angka 1 berarti dipilih dan angka 0 berarti tidak dipilih oleh mahasiswa. Di bawah ini adalah data akhir berjumlah 30 data yang telah di seleksi dan ditransformasi :

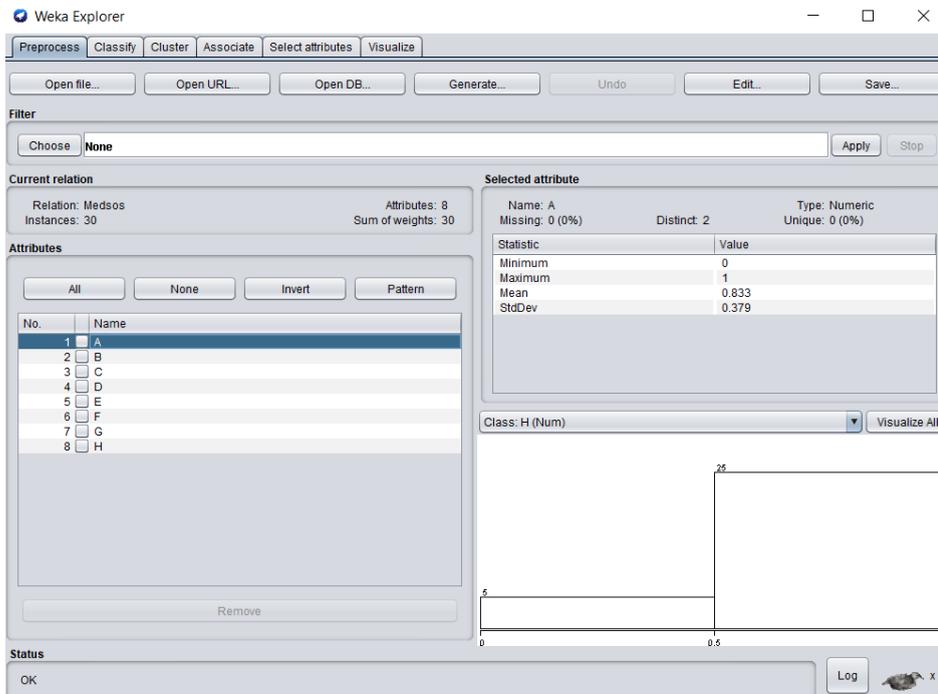
Table 2: Data koesioner yang telah diubah

A	B	C	D	E	F	G	H
1	1	0	1	1	1	1	1
1	0	1	0	1	0	0	0
1	0	0	1	0	0	0	0
1	1	1	0	0	1	0	0
1	1	1	1	0	1	1	0
1	0	1	1	0	0	0	0
1	1	1	1	1	1	1	0
1	0	0	0	0	1	0	1

1	0	0	0	1	0	0	0
1	1	1	1	1	1	1	0
1	1	1	1	1	1	1	0
1	1	0	0	1	0	0	0
1	1	1	1	1	1	1	0
1	0	0	0	1	0	0	0
1	1	1	1	1	1	1	0
0	0	0	0	0	0	0	0
0	0	0	0	1	1	0	0
1	1	0	1	0	1	1	0
1	1	1	0	1	1	0	0
1	1	0	1	0	0	0	0
0	1	1	1	0	1	0	0
1	0	1	0	0	1	0	0
1	1	1	0	0	1	0	0
0	1	0	0	0	1	0	0
0	0	1	1	0	0	0	0
1	0	0	0	0	1	0	0
1	1	0	0	1	1	0	0
1	1	0	1	0	0	0	0
1	1	0	1	0	0	0	0
1	1	1	0	1	1	0	0

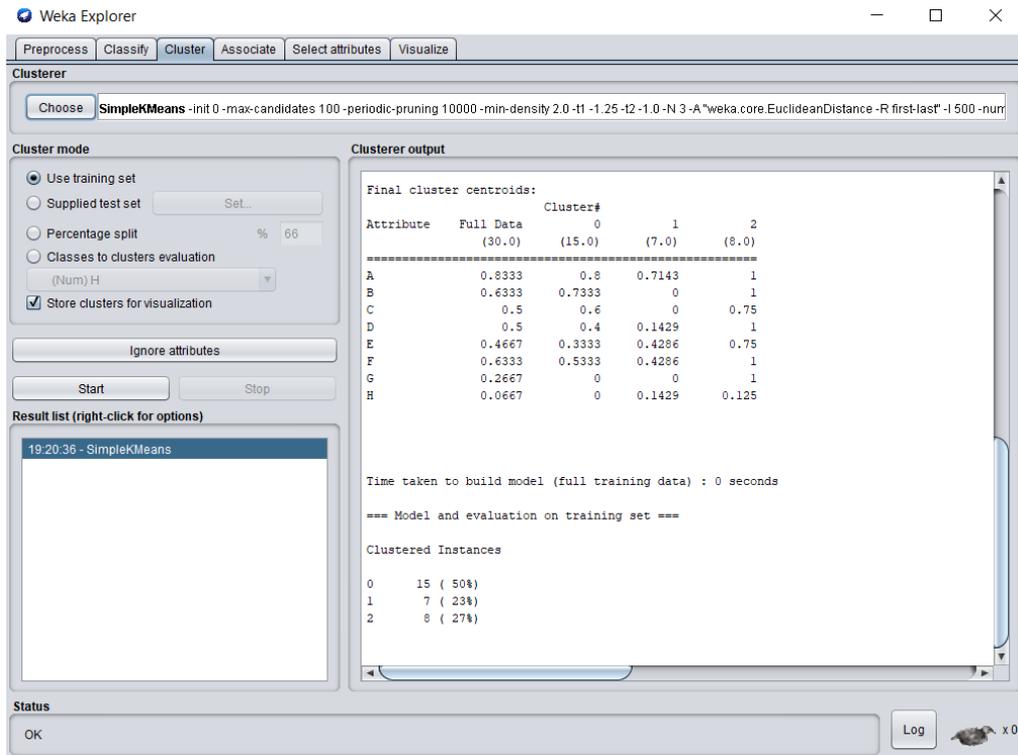
### 3.2 Pengolahan Data

Pada tahap pengolahan, data yang telah di ubah pada Tabel 2 diproses menggunakan aplikasi WEKA.



Gambar 1: Tampilan Data pada WEKA

Selanjutnya setelah data telah terbaca pada aplikasi WEKA maka data diproses kembali menggunakan algoritma clustering simple k means dan menghasilkan seperti pada Gambar 2.



Gambar 2: Hasil Pengolahan data

Pada Gambar 2 merupakan hasil berdasarkan algoritma *clustering simple k-means* yang dimana clusternya dibagi menjadi 3 *cluster* sebagai berikut:

1. Cluster 0 dinamakan Pengguna Pasif, Pada *cluster* ini 50% responden umumnya tidak melakukan kegiatan yang banyak pada media sosial mempunyai hasil mengirim pesan dan menyukai kiriman.
2. Cluster 1 dinamakan Pengguna Normal, Pada *cluster* ini 27% responden tidak jauh berbeda dengan cluster sebelumnya hanya saja responden bertambah satu kegiatan mempunyai hasil mengirim pesan, membaca berita dan bermain game
3. Cluster 2 dinamakan Pengguna Aktif, Pada *cluster* ini 23% responden melakukan banyak kegiatan dalam sosial media dimana mempunyai hasil mengirim pesan, menyukai kiriman, update status, membaca berita dan berbelanja

Dari penjelasan diatas terdapat 3 cluster, dimana Pada cluster 0 termasuk dalam kategori pengguna yang menggunakan media sosial secara normal, lalu pada cluster 1 termasuk dalam kategori pengguna yang menggunakan media sosial secara sedang, dan pada cluster 2 termasuk dalam kategori yang menggunakan media sosial secara aktif.

#### 4 KESIMPULAN

Dari hasil penelitian ini, kami melakukan pengolahan data mining dengan metode clustering dimana algoritma yang digunakan adalah *Simple K-Means*. Tujuan utama penelitian ini yaitu untuk mengelompokan kegiatan yang biasa digunakan oleh mahasiswa pada media sosial. Dari Hasil perhitungan dapat disimpulkan bahwa :

1. Segmentasi dilakukan dengan bantuan tools WEKA untuk perhitungannya.
2. Berdasarkan perhitungan, dihasilkan 3 *cluster* berbeda yang sesuai dengan

ciri masing-masing responden.

3. Diperoleh 50% responden pada *cluster* 0 yang bernama “Pengguna Pasif” yang kegiatannya meliputi mengirim pesan dan menyukai kiriman.

Dari hasil penelitian ini, diperoleh data pengelompokan tipe pengguna media sosial dan kegiatannya. Target pengelompokan tersebut dapat digunakan *software developer* mengembangkan fitur atau kecanggihan aplikasi agar semakin di minati oleh pengguna di kalangan mahasiswa, berdasarkan bahan informasi mengenai fitur apa saja yang sering digunakan.

## Referensi

- Novia Ika Setyani. (2013). Penggunaan Media Sosial Sebagai Sarana Komunikasi Bagi Komunitas.
- Gibran Rabbany, Aripin. (2016). Analisis Aturan Asosiasi Menggunakan Algoritma Apriori Untuk Menentukan Inventori Apotek.
- B. S. d. A. R. B. Tahta Alfina. (2012). Analisa Perbandingan Metode Hierarchical Clustering, K Means Dan Gabungan Keduanya Dalam Cluster Data (studi kasus : problem kerja praktek jurusan teknik industri ITS). Vols. 1, (sept, 2012), no. 2301-9271,p. 2.
- M. N. M. N. S. Ediyanto. (2013). pengklasifikasikan karakteristik dengan metode K-Means cluster analysis. Vols. 02, No.2, hal 133 - 136.
- Agusta, Y. (2007). K-Means - Penerapan, Permasalahan dan Metode Terkait. Jurnal Sistem dan Informatika Vol.3 , 4760.
- Slamet, P, Armadyah, A, Suryanto, M. (2013). ANALISIS KEPUASAN PUBLIK MENGGUNAKAN WEKA DALAM MEWUJUDKAN *GOOD GOVERNANCE* DI KOTA YOGYAKARTA. STMIK AMIKOM Yogyakarta. Vol.14 No 2.