ARTIKEL

IMPLEMENTASI METODE *K-MEANS* UNTUK KEPINDAHAN PENDUDUK STUDI KASUS KANTOR DESA TUGUREJO KABUPATEN KEDIRI

Oleh:

FREDY NUR AZIZY

13.1.02.03.0075

Dibimbing oleh :

1. **M.Rizal Arief, S.T.,M.Kom**
2. **Fajar Rohman Hariri, M.Kom**

TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS NUSANTARA PGRI KEDIRI
2017

Artikel Skripsi
Universitas Nusantara PGRI Kediri

SURAT PERNYATAAN
ARTIKEL SKRIPSI TAHUN 2017

Yang bertanda tangan di bawah ini:

Nama Lengkap	:FREDY NUR AZIZY
NPM	:13.1.03.02.0075
Telepun/HP	:085649111300
Alamat Surel (Email)	:fredynurazizy@gmail.com
Judul Artikel	:Implementasi Metode K-Means untuk Kepindahan Penduduk Studi Kasus Kantor Desa Tugurejo Kabupaten Kediri
Fakultas – Program Studi	:Teknik – Teknik Informatika
Nama Perguruan Tinggi	:Universitas Nusantara PGRI Kediri
Alamat Perguruan Tinggi	:JL. K.H Ahmad Dahlan.76 Kediri

Dengan ini menyatakan bahwa :

a. artikel yang saya tulis merupakan karya saya pribadi (bersama tim penulis) dan bebas plagiarisme;
b. artikel telah diteliti dan disetujui untuk diterbitkan oleh Dosen Pembimbing I dan II.

Demikian surat pernyataan ini saya buat dengan sesungguhnya. Apabila di kemudian hari ditemukan ketidaksesuaian data dengan pernyataan ini dan atau ada tuntutan dari pihak lain, saya bersedia bertanggungjawab dan diproses sesuai dengan ketentuan yang berlaku.

Mengetahui		Kediri, 8 Agustus 2017
Pembimbing I	Pembimbing II	Penulis,
M.Rizal Arief, S.T.,M.Kom 0716027504	Fajar Rohman Hariri, M.Kom 0715058902	FREDY NUR AZIZY 13.1.03.02.0075

FREDY NUR AZIZY | 13.1.03.02.0075
Teknik – Teknik Informatika
simki.unpkediri.ac.id
|| 1||

IMPLEMENTASI METODE *K-MEANS* UNTUK KEPINDAHAN PENDUDUK STUDI KASUS KANTOR DESA TUGUREJO KABUPATEN KEDIRI

FREDY NUR AZIZY
13.1.03.02.0075
Teknik – Teknik Informatika
fredynurazizy@gmail.com
M.Rizal Arief, S.T.,M.kom dan Fajar Rohman Hariri, M.kom
UNIVERSITAS NUSANTARA PGRI KEDIRI

ABSTRAK

FREDY NUR AZIZY : Implementasi Metode *K-Means* Untuk Kepindahan Penduduk Studi Kasus Kantor Desa Tugurejo Kabupaten Kediri.

Kata Kunci : Pengklasifikasian, *Data Mining*, *Clustering*, *K-Means*, *Euclidean Distance*.

Penelitian ini dilatar belakangi hasil pengamatan peneliti, bahwa banyaknya data yang dipunyai sebuah instansi bisa menyebabkan kesulitan dalam mengelola data tersebut untuk kepentingan instansi. Akibatnya banyak membuang waktu dalam pengkasifikasian yang dilakukan.

Permasalahan penelitian ini adalah bagaimana cara merancang sistem untuk pengolahan data kantor dengan menggunakan metode *clustering*?

Di dalam penelitian ini menggunakan pendekatan Penelitian Rekayasa Perangkat Lunak dengan subyek penelitian pelayanan kepindahan penduduk di Kantor Desa Tugurejo. Penelitian ini menggunakan algortima *k-means* dalam mengelompokkan data warga mutasi. Data warga mutasi dikelompokkan berdasarkan jumlah kluster yang ditentukan. Selanjutnya, setiap data/obyek ditempatkan berdasarkan kluster terdekat menggunakan konsep jarak *euclidean distance*. Proses tersebut terus dilakukan sampai pusat kluster tidak berubah lagi.

Hasil penelitian ini adalah (1) Melalui algoritma *K-Means* ini terbukti dapat membantu mengelompokan data warga mutasi dalam jumlah besar berdasarkan kluster yang telah ditentukan secara *random* (acak). (2) Dengan menggunakan metode *Clustering* dengan algoritma *K-Means* terbukti dapat mengelola dan mengelompokkan data warga mutasi secara otomatis.

Berdasarkan hasil penelitian ini, direkomendasikan : (1) Tujuan pokok penggunaan algoritma *k-means* ini adalah untuk mengetahui tingkat kepadatan kepindahan penduduk tiap tahun pada Desa Tugurejo. Oleh sebab itu karyawan Kantor Desa Tugurejo sebagai pengguna aplikasi ini harus mengutamakan proses yang mendukung kelancaran penggunaan dan perawatan aplikasi ini. (2) Karyawan Kantor Desa Tugurejo perlu mengembangkan lagi aplikasi yang telah dibuat oleh peneliti.

I. LATAR BELAKANG

Kantor Desa Tugurejo merupakan salah satu Organisasi Perangkat Desa yang bertanggung jawab dalam mencatat dan memberikan informasi mengenai data kependudukan di Desa Tugurejo. Sampai saat ini Kantor Desa Tugurejo masih menemukan permasalahan dalam memberikan pelayanan pada masyarakat Desa Tugurejo untuk dapat mengelola admistrasi yang baik.

Tugas Pokok dan Fungsi Kantor Desa Tugurejo adalah meningkatkan pelayanan kepada masyarakat. Salah satu bagian yang menghadapi masalah dalam data kependudukan pada Desa adalah Sub Bidang Pendaftaran Penduduk untuk lingkup data Pindah Penduduk dan Datang Penduduk antar Desa, Kecamatan,Kabupaten, Kota dan Provinsi. Sistem yang belum terintegrasi dari tingkat pusat ke tingkat daerah adalah kendala utama yang menyebabkan pengadministrasian yang kurang baik. Hal ini dikarenakan dengan adanya data penduduk yang pindah dan datang yang tidak singkron sehingga proses pendokumentasian akan terganggu baik secara manual maupun secara digital. Maka dari itu perlu adanya suatu sistem yang mampu mengorganisir dan memanajemen data-data tersebut agar dapat meminimalisirkan kekeliruan yang sering terjadi.

Metode yang digunakan dalam sistem pengambil keputusan ini adalah metode *K-Means*. *K-Means* merupakan salah satu metode data *clustering* non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster*/kelompok. Metode ini mempartisi data ke dalam *cluster*/kelompok 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 yang lain. Adapun tujuan dari data clustering ini adalah untuk meminimalisasikan *objective function* yang diatur dalam proses *clustering*, yang pada umumnya berusaha meminimalisasikan variasi di dalam suatu *cluster* dan memaksimalisasikan variasi antar *cluster*.

K-Means merupakan algoritma *clustering* yang berulang-ulang. Algoritma *K-Means* dimulai dengan pemilihan secara acak K, K disini merupakan banyaknya *cluster* yang ingin dibentuk. Kemudian tetapkan nilai-nilai K secara *random*, untuk sementara nilai tersebut menjadi pusat dari

cluster atau biasa disebut dengan *centroid*, mean atau "means". *Jurnal Seminar Nasional Aplikasi Teknologi Informasi 2008 (SNATI 2008)*, Yogyakarta, 21 Juni 2008.

II. METODE

Metode *K-Means* bisa menggunakan ukuran kemiripan untuk mengelompokkan suatu obyek. Kemiripan ini diterjemahkan dalam konsep jarak (Santoso, 2007). Semakin dekat jarak, maka semakin tinggi kemiripannya.

Berikut langkah-langkah algoritma *K-Means Clustering* :

a. Menentukan jumlah K *cluster*.
b. Inisialisasi nilai K pusat *cluster* atau nilai *means* atau *centroid*, yang paling sering digunakan adalah dengan cara*random* (acak).
c. Menempatkan setiap data/obyek ke *cluster* terdekat. Kedekatan dua obyek ditentukan berdasar jarak kedua obyek tersebut. Demikian juga kedekatan suatu data ke *cluster* tertentu ditentukan jarak antara data dengan pusat *cluster*. Dalam tahap ini perlu dihitung jarak tiap data dengan satu *cluster* tertentu akan menentukan suatu data masuk dalam *cluster*. Konsep jarak yang digunakan adalah jarak *Euclidean* (L_2-norm).
d. Menghitung kembali pusat *cluster* dengan keanggotaan *cluster* sekarang. Pusat *cluster* adalah rata-rata dari semua data atau obyek dalam *cluster* tertentu. Jika dikehendaki bisa juga memakai median dari *cluster* tersebut.
e. Menghitung kembali pusat *cluster* yang baru jika pusat tidak berubah lagi, maka proses pengklasteran selesai atau kembali ke langkah yang ketiga sampai pusat klaster tidak berubah lagi.

Hasil klaster dengan metode *K Means Clustering* sangat bergantung pada nilai pusat klaster awal yang diberikan. Pemberian nilai awal yang berbeda bisa menghasilkan hasil klaster yang berbeda. Ada beberapa cara member nilai awal misalnya dengan mengambil sampel awal dari data atau member nilai awal secara *random* (acak).

Beberapa permasalahan yang sering muncul pada saat menggunakan metode *K-Means* untuk melakukan pengelompokkan data adalah :

a. Ditemukannya beberapa model klastering yang berbeda.
b. Pemilihan jumlah klaster yang paling tepat.

c. Kegagalan untuk *converge*.
d. Pendeteksian *outlier*.
e. Bentuk masing-masing klaster.
f. Masalah *overlapping*.

Keenam permasalahan diatas adalah beberapa hal yang perlu di perhatikan pada saat menggunakan *K-Means* dalam mengelompokkan data. Permasalahan 1 umumnya disebabkan oleh perbedaan proses inisialisasi anggota masing-masing klaster. Proses initialisasi yang sering digunakan adalah proses inisialisasi secara *random*.

Permasalahan 2 merupakan masalah laten dalam metode *K-Means*. Beberapa pendekatan telah digunakan dalam menentukan jumlah klaster yang paling tepat untuk suatu dataset yang dianalisa termasuk diantaranya *Partition Entropy* (PE) dan GAP *Statistics*. Metode-metode tersebut tidak sama dengan pendekatan yang digunakan oleh *K-Means* dalam mempartisi data items ke masing-masing klaster.

Permasalahan kegagalan untuk *converge*, secara teori memungkinkan untuk terjadi dalam kedua metode *K-Means* yang dijelaskan di dalam tulisan ini. Kemungkinan ini akan semakin besar terjadi untuk metode *Hard K-Means*, karena setiap data di dalam dataset dialokasikan secara tegas (*hard*) untuk menjadi bagian dari suatu klaster tertentu. Kejadian seperti ini tentu akan mengakibatkan pemodelan tidak akan berhenti dan kegagalan untuk *converge* akan terjadi. Untuk *K-Means*, walaupun ada, kemungkinan permasalahan ini untuk terjadi sangatlah kecil, karena setiap data dilengkapi dengan *membership function* untuk menjadi anggota klaster yang ditemukan.

Permasalahan keempat merupakan permasalahan umum yang terjadi hampir disetiap metode yang melakukan pemodelan terhadap data. Khusus untuk metode *K-Means* hal ini memang menjadi permasalahan yang cukup menentukan. Beberapa hal yang perlu diperhatikan dalam melakukan pendeteksian *outlier* dalam proses pengelompokkan data termasuk bagian menentukan apakah suatu data item merupakan *outlier* dari suatu klaster tertentu dan apakah data dalam jumlah kecil yang membentuk suatu klaster tersendiri dapat dianggap sebagai *outlier*. Proses ini memerlukan suatu pendekatan khusus yang berbeda dengan proses pendeteksian *outlier* di dalam suatu dataset yang hanya terdiri dari satu populasi yang homogen.

Permasalahan kelima adalah menyangkut bentuk klaster yang ditemukan. Tidak seperti metode data *clustering* lainnya termasuk *Mixture*

Modelling, *K-Means* umumnya tidak mengindahkan bentuk dari masing-masing klaster yang mendasari model yang terbentuk, walaupun secara natural masing-masing klaster umumnya berbentuk bundar. Untuk *dataset* yang diperkirakan mempunyai bentuk yang tidak biasa, beberapa pendekatan perlu untuk diterapkan.

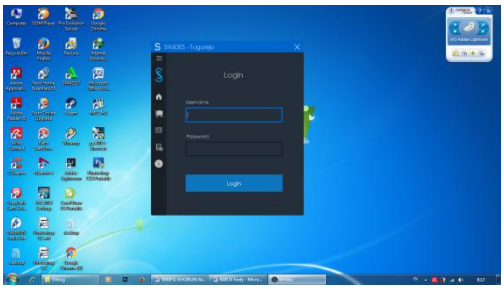
Masalah *over lapping* sebagai permasalahan terakhir sering sekali diabaikan karena umumnya masalah ini sulit terdeteksi. Hal ini terjadi untuk metode *Hard K-Means* dan *Fuzzy K-Means*, karena secara teori metode ini tidak diperlengkapi *feature* untuk mendeteksi apakah didalam suatu klaster ada klaster lain yang kemungkinan tersembunyi. (Yunus Ariadi Saputra, 2014)

III. HASIL DAN KESIMPULAN

1. Hasil

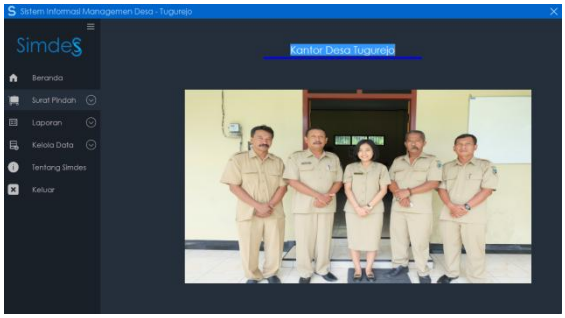
a. Halaman login

Login digunakan untuk masuk kedalam aplikasi Sistem Kepindahan Penduduk Desa.



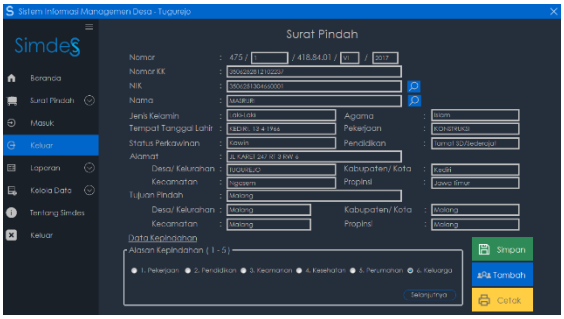
b. Halaman utama

Halaman ini mencakup semua halaman form yang ada pada sistem seperti surat pindah, laporan pindah, perhitungan dan kelola data.



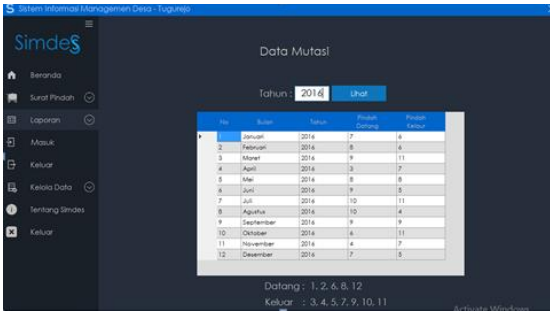
c. Input Surat Pindah

Halaman ini digunakan oleh *administrator* untuk melakukan pelayanan dan pengolahan data kepindahan penduduk warga. Simdes berisi : Surat Pindah (Masuk dan Keluar), Laporan Pindah (Masuk dan Keluar), Kelola data Warga Desa Tugurejo (Data penduduk dan Kepala Keluarga), serta menu Tentang Desa Tugurejo.



d. *Hasil output Perhitungan*

Halaman ini berisi outputan hasil proses perhitungan metode *K-Means*.



2. Kesimpulan

Berdasarkan hasil yang telah disajikan pada bab IV dan V maka :

1. Telah dihasilkan rancangan Sistem Kepindahan Penduduk Desa yang berisi proses atau tahap yang harus lakukan sebelum menyusun program aplikasi sistem Kepindahan Penduduk.
2. Telah dihasilkan program aplikasi Sistem Kepindahan Penduduk Desa yang dapat dijalankan oleh admin yang bekerja di Kantor Desa Tugurejo.

IV. DAFTAR PUSTAKA

Ariadi, Yunus S. 2014. Analisis Data Mahasiswa Menggunakan Algoritma *K-Means Clustering* (Studi Kasus Prodi Teknik Informatika Fakultas Teknik UNP Kediri). Universitas Nusantara PGRI Kediri

Didik Setiawan, Yhoni Agus Setya Mahendra. 2015. *Perancangan Sistem Informasi Penduduk Pada Kantor Desa Kebonsari*. Program Studi Teknik Informatika Universitas Yudharta Pasuruan Diunduh Pada 21 Desember 2016

Diyah Prihastini. 2016. *Sistem informasi pelayanan kependudukan*. Universitas Negeri Semarang : diunduh pada 21 Desember 2016

Dwi Noviati Nango. 2012. *Penerapan Algoritma K-Means Untuk Clustering Data Anggaran Pendapatan Belanja Daerah Di Kabupaten Xyz.* Universitas Negeri Gorontalo diunduh pada 11 Nopember 2016

Irwan Budiman, M.Kom.. 2012. *Data clustering menggunakan metodologi crisp-dm untuk pengenalan pola proporsi pelaksanaan tridharma*. Universitas Diponegoro

Kadir, Abdul. 2013. *Pengenalan Sistem Informasi*. Yogyakarta : Andi

Moh.Shobiri. 2013. *Profil Desa Tugurejo*. Kantor Desa Tugurejo

Rina Puspitasari. 2013. *Aplikasi pembayaran iuran sekolah dan buku lembar kerja siswa pada sekolah menengah atas negeri 1 punung*. Universitas Surakarta : diunduh pada 26 Desember 2016

Tedy Rismawan, Sri Kusumadewi. 2008. *Aplikasi K-Means Untuk Pengelompokkan Mahasiswa Berdasarkan Nilai Body Mass Index (Bmi) & Ukuran Kerangka*. Jurusan Teknik Informatika, FTI, UII diunduh pada 11 Nopember 2016

Vivin Dwi Retnaningsih. 2012. *Pengklasifikasian Data Sekolah Pengguna Internet Pendidikan Menggunakan Teknik Clustering Dengan Algoritma K-Means Studi Kasus Pt Telkom Surabaya*. Universitas Nusantara PGRI Kediri