

CLUSTERING PRODUKSI BARANG PADA LALU LINTAS ANGKUTAN UDARA MENGUNAKAN ALGORITMA K-MEANS

CLUSTERING OF CARGO PRODUCTION ON THE AIR TRANSPORT TRAFFIC USING K- MEANS ALGORITHM

Deyan Havith Dailamy¹, Dr. Purba Daru Kusuma², Anggunmeka Luhur Prasasti³

^{1,2,3}Prodi S1 Sistem Komputer, Fakultas Teknik Elektro, Universitas Telkom

¹deyanhavith@students.telkomuniversity.ac.id, ²purbadarukusuma@telkomuniversity.ac.id,

³anggunmeka@telkomuniversity.ac.id

Abstrak

Setiap hari bandara yang berada di Indonesia melakukan rute penerbangan baik dalam negeri dan luar negeri dimana setiap bandara memiliki lalu lintas angkutan seperti lalu lintas penumpang, barang, dan bagasi dengan jumlah yang sangat besar dan berubah-ubah setiap tahunnya. Maka diperlukan pemanfaatan data dengan cara membuat sistem Data Mining dari data tersebut yang akan digunakan untuk kepentingan tertentu, salah satunya adalah melakukan clustering atau pengelompokkan dengan menggunakan metode clustering yang merupakan metode penganalisaan data yang tujuannya adalah untuk mengelompokkan data dengan karakteristik yang sama ke suatu wilayah yang sama dan dengan karakteristik yang berbeda ke wilayah yang lain. Sedangkan untuk algoritma yang digunakan adalah algoritma K-Means yang merupakan salah satu algoritma terbaik dan paling populer dalam metode Clustering dimana K-Means mencari partisi yang optimal dari data dengan meminimalkan kriteria jumlah kesalahan kuadrat dengan prosedur iterasi yang optimal menggunakan beberapa variabel yaitu kedatangan barang (bongkar) dan keberangkatan barang (muat) di suatu penerbangan dalam negeri dan luar negeri di Indonesia.

Kata kunci : *clustering*, algoritma *K-Means*, klasifikasi, lalu lintas angkutan udara.

Abstract

Every day the airports in Indonesia do flight routes both domestically and abroad where each airport has transportation traffic such as passenger, goods, and baggage traffic with a very large amount and changes every year. So it is necessary to use data by means of make a Data Mining system from the data that will be used for certain purposes, one of which is to do clustering or grouping using clustering method which is a method of analyzing data whose purpose is to group data with the same characteristics into the same area and with different characteristics to another area. Whereas the algorithm used is the K-Means algorithm which is one of the best and most popular algorithms in the Clustering method where K-Means searches for the optimal partition of data by minimizing the criteria for the number of squared errors with an optimal iteration procedure using several variables, namely the arrival of goods (unloading and departure of goods (loading) on a domestic and foreign flight in Indonesia.

Keywords: clustering, K-Means algorithm, classification, air freight traffic.

1. Pendahuluan

Transportasi mempunyai peranan yang sangat penting dan strategis dalam mendukung, mendorong dan menunjang segala aspek kehidupan [1]. Transportasi udara menjadi kian sangat penting karena luasnya wilayah Negara Kesatuan Republik Indonesia yang dipisahkan oleh perairan karena dengan menggunakan transportasi udara ini dapat menghubungkan wilayah-wilayah tersebut dengan waktu tempuh yang cepat[2]. Dalam menjalankan transportasi udara, Bandara Udara merupakan komponen penting di dalamnya. Dimana Bandara Udara merupakan kawasan di daratan atau perairan dengan batas-batas tertentu yang digunakan sebagai tempat pesawat udara mendarat dan lepas landas, naik turun penumpang, bongkar muat barang, dan tempat perpindahan intra dan antarmoda transportasi, yang dilengkapi dengan fasilitas keselamatan dan keamanan penerbangan, serta fasilitas pokok dan fasilitas penunjang lainnya.

Setiap penerbangan dalam negeri dan luar negeri di suatu bandara tentunya memiliki lalu lintas angkutan udara seperti lalu lintas pesawat, lalu lintas penumpang, lalu lintas bagasi, dan lalu lintas barang. Dimana lalu lintas angkutan udara yang disebutkan diatas diarsipkan dalam sebuah data setiap tahunnya dari setiap bandara di Indonesia oleh Badan Pusat Statistik Indonesia. Data tersebut terdiri dari data barang yang dibongkar dan dimuat setiap tahunnya yang diperoleh dari setiap bandara di Indonesia. Data tersebut berjumlah sangat banyak dan bersifat kuantitatif yaitu jenis data yang dapat dihitung, berupa angka atau nominal[4] dan fluktuatif yaitu keadaan dimana data yang didapatkan cenderung tidak stabil dan berubah-ubah setiap tahunnya. Hal ini dapat menyulitkan bidang penerbangan khususnya

pihak pengelola bandara dalam melakukan perkembangan sumber daya manusia dan perencanaan pembangunan sarana maupun prasarana yang berkelanjutan di masa datang.

Pemilihan Algoritma K-Means sebagai algoritma yang digunakan karena memiliki kelebihan dibandingkan algoritma lainnya. Kelebihan pada algoritma K-means yaitu [6]: Mudah untuk diimplementasikan dan dijalankan, waktu yang dibutuhkan untuk menjalankan pembelajaran ini relatif cepat, mudah untuk diadaptasi, dan umum digunakan. Hasil dari penelitian tugas akhir ini adalah dapat membuat visualisasi pengelompokan bandara di Indonesia berdasarkan provinsi yang dikhususkan pada produksi barang yang memiliki tingkat arus lalu lintas angkutan udara yang sangat tinggi, tinggi, sedang, rendah, dan dan sangat rendah. sehingga nantinya akan ada lima cluster yang dibuat didalam aplikasi ini, diharapkan aplikasi ini dapat digunakan untuk perencanaan sumber daya manusia dan sarana bandar udara di masa yang akan datang.

2. Dasar Teori

2.1. Publikasi statistik angkutan udara

Publikasi statistik angkutan udara adalah suatu publikasi yang diterbitkan setiap tahun oleh Badan Pusat Statistik yang bertujuan untuk memberikan informasi kepada pengguna data, baik instansi pemerintah maupun swasta mengenai aktivitas penerbangan baik domestik maupun internasional seluruh bandara di Indonesia serta perkembangannya dalam beberapa tahun [13].

Pengumpulan data angkutan udara tahun 2016 terdiri atas beberapa kegiatan sesuai jenis data yang dikumpulkan. Jenis dan sumber data diuraikan sebagai berikut:

a). Registrasi Pesawat

Data yang dikumpulkan berkaitan dengan registrasi pesawat mencakup berbagai tipe pesawat yang dimiliki oleh seluruh operator penerbangan yang ada di Indonesia. Sumber data registrasi pesawat berasal dari Direktorat Jenderal Perhubungan Udara, Kementerian Perhubungan Republik Indonesia.

b). Produksi Perusahaan Penerbangan

Data produksi meliputi kmpesawat dan keberangkatan pesawat, jam terbang pesawat, penumpang dan barang yang diangkut, kmpenumpang, faktor muatan penumpang, dan faktor muatan barang. Sumber data produksi perusahaan penerbangan berasal dari Direktorat Jenderal Perhubungan Udara, Kementerian Perhubungan.

c). Lalu Lintas Angkutan Udara di Bandar Udara

Data lalu lintas angkutan udara terdiri atas lalu lintas pesawat, penumpang, bagasi, barang, dan pos/paket [13].

1. 2.2. K-Means Clustering

Clustering adalah proses mengelompokkan atau penggolongan objek berdasarkan informasi yang diperoleh dari data yang menjelaskan hubungan antar objek dengan prinsip untuk memaksimalkan kesamaan antar anggota satu kelas dan meminimumkan kesamaan antar kelas/cluster [12]. Clustering membagi data ke dalam grup-grup yang mempunyai obyek yang karakteristiknya sama.

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 di dalam suatu cluster dan memaksimalkan variasi dengan data yang ada di cluster lainnya [12]. Menurut Sarwono [12], Berikut adalah langkah-langkah dari algoritma K-Means:

1. Menentukan banyak *k-cluster* yang ingin dibentuk.
2. Membangkitkan nilai random untuk pusat cluster awal (*centroid*) sebanyak *k-cluster*.
3. Menghitung jarak setiap data input terhadap masing-masing centroid menggunakan rumus jarak *Euclidian (Euclidian Distance)* hingga ditemukan jarak yang paling dekat dari setiap data dengan centroid.
4. Mengklasifikasikan setiap data berdasarkan kedekatannya dengan centroid (jarak terkecil).

Berikut rumus Euclidian:

$$d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

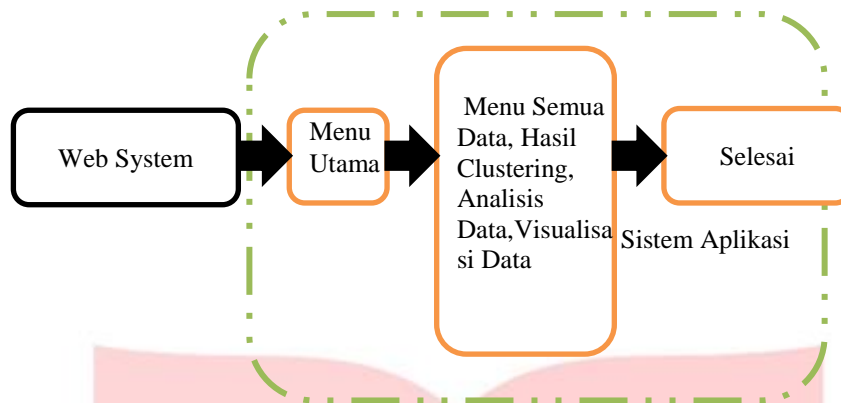
dengan d x_i , y_i adalah jarak antara cluster x dengan pusat cluster y pada kata ke i , x_i adalah bobot kata ke i pada cluster yang ingin dicari jaraknya, y_i bobot kata ke i pada pusat cluster

5. Mengupdate nilai centroid. Nilai *centroid* baru diperoleh dari rata-rata cluster yang bersangkutan.
6. Melakukan perulangan dari langkah 2 hingga 5 hingga anggota tiap cluster tidak ada yang berubah.
7. Jika langkah 6 telah terpenuhi, maka nilai rata-rata pusat *cluster* (μ_j) pada iterasi terakhir akan digunakan sebagai parameter untuk menentukan klasifikasi data.

3. Pembahasan

3.1. Gambaran umum sistem

Berdasarkan permasalahan yang telah didapatkan beserta dasar dari berbagai sumber, maka pada bagian ini merupakan konsep penyelesaian masalah dengan melakukan perancangan dan implementasi. Data yang digunakan pada sistem yang akan dibuat adalah berasal dari Badan Pusat Statistik Indonesia mengenai Statistik Transportasi Udara 2016. Adapun gambaran umum dari sistem ini terdapat pada Gambar 3.1:

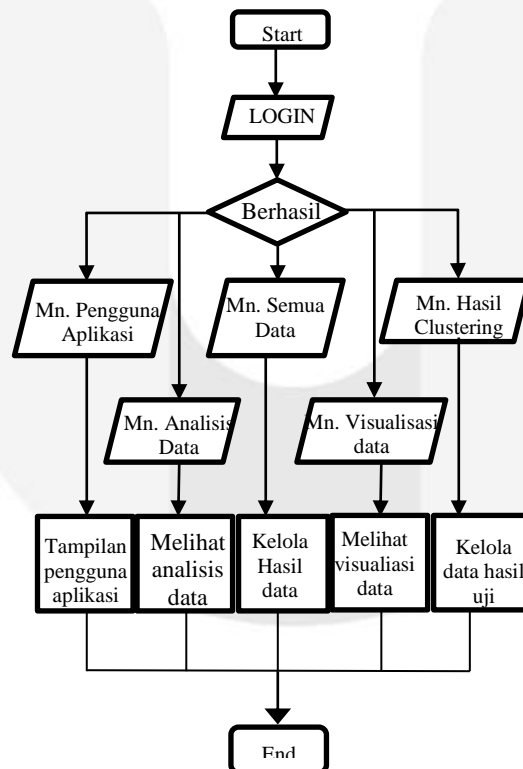


Gambar 3.1. Perancangan Aplikasi

Web System berfungsi sebagai *interface* aplikasi. Setelah masuk di aplikasi maka akan menampilkan menu utama dari aplikasi. Setelah itu bisa memilih menu yang diinginkan.

3.2. Rancangan Sistem

Pada gambar 3.2 dibawah ini menunjukkan gambar rancangan aplikasi *Clustering* Produksi Barang Pada Lalu Lintas Angkutan Udara Menggunakan Algoritma K-Means. Proses dimulai dengan melakukan login, apabila *login* tidak berhasil maka sistem akan kembali pada halaman awal login. Apabila *login* berhasil maka sistem akan menampilkan menu.



Gambar 3.2. Gambaran Umum Rancangan Aplikasi

2. 3.3. Unified Modeling Language (UML)

Aplikasi dibangun dengan menggunakan *Unified Modeling Language* (UML). UML merupakan bahasa visual untuk pemodelan dan komunikasi mengenai sebuah sistem dengan menggunakan diagram yang terdiri dari *Use Case Diagram*, *Activity Diagram* dan *Class Diagram*.

4. Implementasi dan Pengujian Sistem

Tahap implementasi sistem akan dilakukan setelah tahap perancangan sistem selesai. Pada tahapan ini hasil rancangan diimplementasikan pada Bahasa pemrograman yang bertujuan untuk membuat model-model perancangan, sehingga sistem siap digunakan.

3. 4.1. Pengujian dan Implementasi Aplikasi

Pada sub bab ini merupakan hasil dari UML yang telah diaplikasikan kedalam aplikasi sehingga diharapkan dengan dibuatnya web dapat mempermudah perhitungan cluster oleh pengguna yang bersangkutan.

4.1.1. Halaman Penggunaan Aplikasi

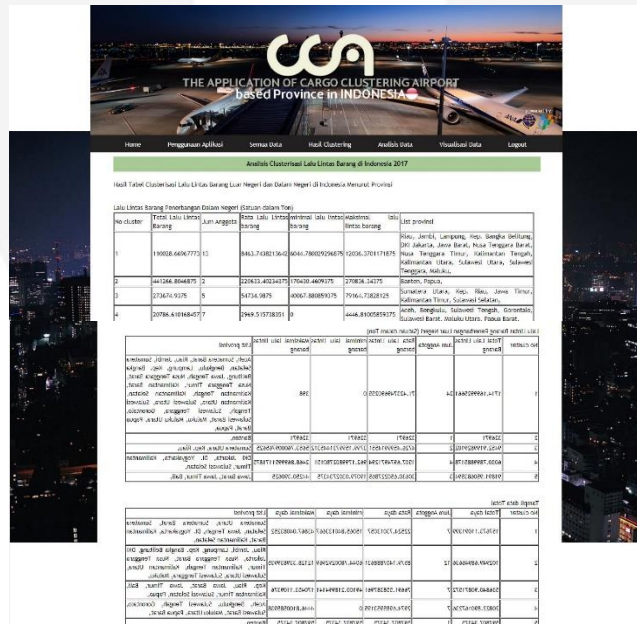
Pada gambar 4.3 berisi tentang prosedur dan tata cara penggunaan menjalankan aplikasi ini. Pada halaman ini berguna untuk memudahkan user saat akan melakukan eksekusi pada sistem.



Gambar 4.1. Halaman Penggunaan Aplikasi

4.1.2. Halaman Analisis Data

Pada halaman ini merupakan halaman Analisis Data dimana data-data yang sudah dikelompokkan ke dalam cluster dimasukkan kedalam tabel agar mudah dalam melakukan analisis. Hasil Analisis Data ini terbagi menjadi dua tabel yaitu lalu lintas dalam negeri dan luar negeri. Tampilan halaman Analisis Data dapat dilihat pada gambar 4.6. dibawah ini.



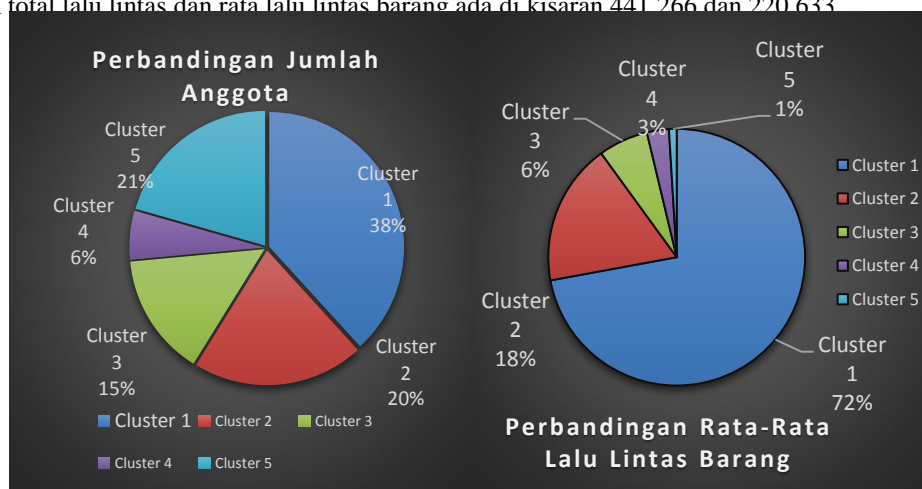
Gambar 4.2. Halaman Analisis Data.

4. 4.2. Hasil Analisis Pengelompokkan Lalu Lintas Barang berdasarkan Cluster

Tabel 4.1. Hasil Analisis Data Dalam Negeri

Cluster	Total Barang	Jumlah	Rata Lalu Lintas barang	Minimal lalu lintas barang	Maksimal lalu lintas barang	List provinsi
1	441266.8	2	220633.4	170430.4	270836.3	Banten, Papua,
2	273674.9	5	54734.9	40067.8	79164.7	Sumatera Utara, Kep. Riau, Jawa Timur, Kalimantan Timur, Sulawesi Selatan,
3	134276	7	19182.2	14791.3	23324.3	Sumatera Barat, Sumatera Selatan, Jawa Tengah, DI. Yogyakarta, Bali, Kalimantan Barat, Kalimantan Selatan,
4	110028.6	13	8463.7	6044.7	12036.3	Riau, Jambi, Lampung, Kep. Bangka Belitung, DKI Jakarta, Jawa Barat, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Tengah, Kalimantan Utara, Sulawesi Utara, Sulawesi Tenggara, Maluku,
5	20786.6	7	2969.5	0	4446.8	Aceh, Bengkulu, Sulawesi Tengah, Gorontalo, Sulawesi Barat, Maluku Utara, Papua Barat,

Pada tabel 4.1. dapat dianalisis bahwa untuk lalu lintas barang/kargo dengan menggunakan jasa transportasi udara dalam negeri antar provinsi masih terbilang kecil karena sesuai dengan tabel dan gambar 4.9 diatas ada sekitar 13 provinsi atau 36% yang termasuk dalam cluster 4 ini yaitu Riau, Jambi, Lampung, Kep. Bangka Belitung, DKI Jakarta, Jawa Barat, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Tengah, Kalimantan Utara, Sulawesi Utara, Sulawesi Tenggara, dan Maluku dengan total lalu lintas barang hanya 110.028 ton dan rata lalu lintas barang hanya 8.463 ton. Cluster 4 ini hanya sekitar 3% dari jumlah total lalu lintas pada Cluster 1 yang dimiliki oleh Aceh dan Banten dengan total lalu lintas dan rata lalu lintas barang ada di kisaran 441.266 dan 220.633



Gambar 4.3. Diagram Analisis Data Dalam Negeri.

Selanjutnya terdapat 7 provinsi yang masuk pada cluster 5 yaitu lalu lintas barang sangat kecil karena hanya berkisar 2.969 ton (Aceh, Bengkulu, Sulawesi Tengah, Gorontalo, Sulawesi Barat, Maluku Utara, Papua Barat). sehingga total ada 20 provinsi yang bertempat di cluster terendah yaitu cluster 4 dan 5 (kecil dan sangat kecil). Hanya berkisar 1% dari jumlah total lalu lintas barang di Cluster 1.

Sedangkan untuk lalu lintas barang yang sangat tinggi atau Cluster 1 di Indonesia dalam negeri hanya dimiliki 2 provinsi saja yaitu Banten dan Papua dengan total lalu lintas barang 441.266 ton dimana memegang 72% dalam jumlah total lalu lintas barang. Cluster selanjutnya atau cluster 2 yaitu lalu lintas barang yang tinggi (273.674 ton total lalu lintas barang) dengan beranggotakan 5 provinsi. Tetapi jarak antara cluster sangat tinggi atau Cluster 1 dengan yang cluster yang tinggi atau Cluster 2 sangat jauh yaitu sekitar 167.592 ton (441266 ton-273674 ton) atau setara 18% saja. Jika dilihat bahwa mengapa Cluster 1 merupakan Cluster yang memiliki tingkat total lalu lintas barang yang tinggi dibandingkan Cluster lainnya karena Cluster ini diisi oleh bandara terbesar di Indonesia yaitu Bandara Soekarno-Hatta dengan jumlah pesawat dan jadwal penerbangan yang banyak.

Sedangkan untuk lalu lintas barang yang sangat tinggi atau Cluster 1 di Indonesia dalam negeri hanya dimiliki 2 provinsi saja yaitu Banten dan Papua dengan total lalu lintas barang 441.266 ton dimana memegang 72% dalam

jumlah total lalu lintas barang .Cluster selanjutnya atau cluster 2 yaitu lalu lintas barang yang tinggi (273.674 ton total lalu lintas barang) dengan beranggotakan 5 provinsi. Tetapi jarak antara cluster sangat tinggi atau Cluster 1 dengan yang cluster yang tinggi atau Cluster 2 sangat jauh yaitu sekitar 167.592 ton (441266 ton-273674 ton) atau setara 18% saja. Jika dilihat bahwa mengapa Cluster 1 merupakan Cluster yang memiliki tingkat total lalu lintas barang yang tinggi dibandingkan Cluster lainnya karena Cluster ini diisi oleh bandara terbesar di Indonesia yaitu Bandara Soekarno-Hatta dengan jumlah pesawat dan jadwal penerbangan yang banyak.

Tabel 4.2. Hasil Analisis Data Luar Negeri

Cluster	Total L. Lintas Barang	Jum Anggota	Rata Lalu Lintas barang	minimal lalu lintas barang	Maksimal lalu lintas barang	List provinsi
1	326971	1	326971	326971	326971	Banten,
2	91891.9	3	30630.6	19079	44250.3	Jawa Barat, Jawa Timur, Bali,
3	9452.9	2	4726.4	3799.1	5653.7	Sumatera Utara, Kep. Riau,
4	6030.7	4	1507.6	962.1	2468.8	DKI Jakarta, DI. Yogyakarta, Kalimantan Timur, Sulawesi Selatan,
5	1714.1	24	71.4	0	398	Aceh, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kep. Bangka Belitung, Jawa Tengah, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, Papua,

Analisis yang bisa didapatkan pada tabel diatas adalah kembali lalu lintas barang luar negeri di Indonesia juga masih terbilang sangat kecil karena ada sekitar 24 provinsi atau 73% yang masuk kedalam Cluster 5 dimana total lalu lintas barang bila dibandingkan dengan Cluster 1 hanya 1 % saja. Hal ini bisa menjadikan analisis kedepan mengapa lalu lintas barang untuk luar negeri terbilang sangat kecil. Salah satu faktor yang bisa mempengaruhi mengapa lalu lintas barang luar negeri dikategorikan sangat rendah adalah daya tarik masyarakat/pelaku usaha itu sendiri yang masih belum menjadikan transportasi udara sebagai sarana kirim barang yang utama. Hal menarik yang ditemukan pada analisis ini yaitu hanya provinsi Jawa Tengah yang merupakan provinsi di Jawa yang masuk ke dalam Cluster 5 sedangkan untuk Provinsi Jawa lainnya masuk kedalam Cluster 1,2, dan 4.

5. Kesimpulan dan Saran

5.1. Kesimpulan

Kesimpulan yang dapat diambil adalah telah diterapkan aplikasi dengan konsep Data Mining menggunakan Algoritma K-Means guna mengelompokkan data lalu lintas angkutan barang menggunakan pesawat terbang di Indonesia baik lalu lintas dalam dan luar negeri. Hasil pengelompokkan digunakan untuk mengetahui provinsi mana yang memiliki tingkat lalu lintas barang yang sangat tinggi hingga sangat rendah sehingga didapatkan beberapa analisis. Hasil analisis-analisis yang telah didapatkan tersebut digunakan untuk memberi saran dan pertimbangan kepada pelaku usaha atau pengelola bandara, khususnya dalam menentukan strategi perencanaan di suatu bandara-bandara di provinsi meliputi sarana, prasarana dan sumber daya manusia yang akan dibutuhkan dimasa depan. Dengan terpenuhi sarana, prasarana, dan sumber daya manusia diharapkan dapat menjadikan bandara udara tersebut andal, efisien, dan memiliki daya saing global sesuai isi UU tentang Tata Nelayan Kebandarudaraan Nasional Pasal 2.

5.2. Saran

Penulis menyadari bahwa aplikasi ini belum sepenuhnya sempurna, masih banyak hal yang dapat dilakukan untuk mengembangkan aplikasi ini agar menjadi lebih baik lagi, diantaranya:

1. Dapat dilakukan penelitian lebih lanjut dengan mengembangkan desain antarmuka sistem, sehingga sistem dapat lebih mudah digunakan.
2. Dapat dilakukan pengembangan lebih lanjut dalam mengkluster lalu lintas barang yaitu berdasarkan masing-masing bandara secara detail seperti dimasukkan data luas dan kapasitas bandara.
3. Dengan berkembangnya teknologi mobile saat ini, diharapkan dapat dilakukan pengembangan lebih lanjut untuk akses sistem ini melalui smartphone dengan berbagai platform yang ada.



Daftar Pustaka

- [1] Manheim, M. L. 1979. *Fundamentals of Transportation System Analysis. Volume I: Basic Concepts*. Cambridge: The MIT Press.
- [2] Jou, R. C., & Chen, T. Y. 2014. Factors affecting public transportation, car, and motorcycle usage. *Transportation Research Part A: The united Media*.
- [3] Sarwono, J., & Prihartono, K. (2012). *Penerbangan Indonesia: Domestik dan Internasional*. Jakarta: PT Elex Media Komputindo.
- [4] Chee, W. L., & Fernandez, J. L. 2013. Factors that Influence the Choice of Mode of Transport in Penang: A Preliminary Analysis.
- [5] Ginting, Selvia Lorena. 2014. Analisis dan Penerapan Algoritma K-Means dalam Data Mining untuk Memprediksikan Masa Studi Mahasiswa Berdasarkan Data Nilai Akademik. Bandung: SNAST 2014.
- [6] Santosa, B. 2007. *Data mining Terapan: Klustering dan Implementasinya*. Surabaya: Graha Ilmu.
- [7] Maimon, O., & Rokach, L. (2010). *Data Mining and Knowledge Discovery Handbook*. London: Springer Science+Business Media.
- [8] Puranti, N,D. 2014. Analisis Data Lalu Lintas Angkutan Barang dengan Maximum Entropy dan Support Vector Machine. Jakarta: IJCC Vol 9.
- [9] Nugroho, Adi. 2004. *Konsep Pengembangan Sistem Basis Data*. Bandung: Informatika.
- [10] Susanti. 2010. *Klasifikasi Kredit Menggunakan Metode Algoritma K-Means Pada Nasabah Pada BPR BKK*. Semarang: Fakultas Ilmu Komputer
- [11] Santosa, Budi. 2007. *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- [12] Astuti, Yuli. 2016. *Clustering Postingan di Twitter Cuaca Provinsi DIY Menggunakan Algoritma K-Means Untuk Informasi Pada Web Pariwisata*. Yogyakarta: STMIK AMIKOM
- [13] Badan Pusat Statistik Indonesia. 2016. *Statistik Transportasi Udara 2016*. Jakarta: BPS-Statistic Indonesia.
- [14] Rusdah. (2011). Analisa Dan Rancangan Sistem Informasi Persediaan Bagasi pada Kereta Api di Indonesia Menggunakan Python. Studi Kasus Ilmu Komunikasi. *Jurnal TELAMATIKA M.KOM*, 3. No. 2. pp. 51-59. Jakarta.
- [15] Chen, S., & Zegras, C. 2016. Rail Transit Ridership. *Transportation Research Record: Journal of the Transportation Using Jupyter Tools*. Research Board.
- [16] Republik Indonesia. 2002. *UU Tatanan Kebandaraan di Indonesia : SE-25/PJ.10/2002* . Jakarta: Direktorat Jenderal Perhubungan Udara.