

LAMPIRAN D
PENGELOMPOKKAN DATA BATUBARA DI INDONESIA
MENGGUNAKAN K- MEANS CLUSTERING

GROUPING OF COAL DATA IN INDONESIA USING K-MEANS
CLUSTERING METHOD

Cut Aisyah Ilmi₁

₁Program S1 Teknik Komputer, Fakultas Teknik Elektro, Universitas Telkom
cutaisyah0808@gmail.ac.id

Abstrak

Tugas akhir ini membahas tentang pembuatan aplikasi *clustering* batubara menggunakan *K-means clustering* yang bertujuan untuk mengelompokkan data batubara (salah satu sektor di Indonesia yang menguntungkan Negara) sehingga terlihat pola dan nilai pada masing-masing *cluster* dan bisa dianalisa dengan mudah informasi yang didapatkan dari pengelompokkan data tersebut.

Kata Kunci : *clustering*, algoritma k-means, batubara

Abstract

This final project discusses the making of coal clustering applications using K-means clustering which aims to classify coal data (one of the sectors in Indonesia that benefits the State) so that patterns and values can be seen in each cluster and can be easily analyzed for informations that are obtained from the data.

Keywords: clustering, k-means algorithm, coal.

1. Pendahuluan

1.1 Latar Belakang

Batubara adalah sedimen organik *hydro-carbon* padat yang terbentuk dari tumbuh-tumbuhan dalam lingkungan bebas oksigen dan terkena pengaruh temperatur serta tekanan yang berlangsung sangat lama. [1] Batubara ditambang karena memiliki banyak manfaat, antara lain sumber tenaga pembangkit listrik, industri produk baja, semen dan juga bahan kimia. Secara ekonomis, batubara memiliki nilai jual yang tinggi karena banyaknya bidang yang membutuhkan bahan bakar fosil tersebut. Di Indonesia sendiri, batubara adalah salah satu sumber daya alam yang paling banyak di *export* dan berada di peringkat 10 besar negara produsen dan eksotir [2].

Di Indonesia, aktivitas pertambangan batubara banyak jumlahnya. Banyaknya aktivitas tersebut menghasilkan data-data yang berkaitan dengan aktivitas pertambangan menjadi banyak juga akan tetapi data-data tersebut masih bersifat mentah dan cukup sulit dianalisis sehingga diperlukan suatu sistem untuk mengolah data-data mentah itu menjadi informasi yang berguna, mudah dipahami dan bersifat otomatis agar bisa berguna dalam sektor pertambangan batubara di Indonesia.

Untuk menangani permasalahan tersebut, penulis ingin membuat aplikasi *clustering* berbasis *web* dimana data mentah yang dimasukkan bisa dianalisis dan didapatkan informasi berharganya secara otomatis melalui *clustering* dengan algoritma k-means. Banyak kasus yang menggunakan k-means *clustering* untuk penganalisaan dan pengelompokan data contohnya pengelompokan mahasiswa menggunakan algoritma k-means [3], pengelompokan data penjualan batubara pada PT. Global Bangkit Utama [4] dan pengelompokan pembangkit listrik di Indonesia menggunakan algoritma k-means [5]. Hasil dari *cluster* beserta analisisnya akan ditampilkan di aplikasi. Dari hasil analisis, kita bisa mendapatkan pola dari cluster data tersebut sehingga memudahkan pengguna untuk menganalisis data.

1.2 Tujuan Masalah

Tujuan masalah pada penelitian tugas akhir adalah

1. Cara membuat aplikasi *clustering* batubara menggunakan metode K-Means *clustering*.
2. Cara menampilkan visualisasi hasil *clustering* data.
3. Cara menganalisa hasil dari K-Means *clustering*.

2. Dasar Teori /Material dan Metodologi/perancangan

2.1 Clustering

Clustering adalah pengelompokan objek ke dalam satu kelas berdasarkan kesamaan karakteristik objek [6].

2.2 Algoritma K-Means

Algoritma ini mengelompokkan sendiri data-datanya tanpa tau target kelasnya terlebih dahulu. Masukannya berupa k jumlah kelompok yang diinginkan. [7].

3. Pembahasan

3.1 Gambaran Sistem Umum

Gambaran sistem umum pada sistem adalah penggambaran secara umum bagaimana sistem tersebut bekerja dan dapat digunakan. Gambaran umum sistem aplikasi ini dimulai dari pengguna yang harus mendaftar terlebih dahulu, kemudian *login* agar bisa memakai aplikasi, setelahnya pengguna bisa mengakses tabel data, visualisasi *clustering*, dan analisis dari hasil *clustering*.

3.2 Analisis Kebutuhan Sistem

3.2.1 Analisis Kebutuhan Data

Data yang digunakan sebagai data mentah pada aplikasi ini adalah data dari kementerian ESDM [8]

3.2.2 Analisis Kebutuhan Perangkat Lunak

Dalam pembangunan aplikasi ini, perangkat lunak yang digunakan adalah R,R studio, SQLite dan SQLite Studio

3.2.3 Analisis Kebutuhan Perangkat Keras

Perangkat keras yang dibutuhkan dalam pembangunan aplikasi ini adalah laptop dengan spesifikasi laptop sebagai berikut : Processor = intel(R) Core(TM) i3-4005U @ 1.70 GHz (4 CPUs), RAM : 4,096 GB dan sistem operasi : windows 10 Pro 64-bit(10.0, Build 17134)

3.2.4 Analisis Kebutuhan Pengguna

Analisis kebutuhan pengguna merupakan fungsi dan fitur yang dikerjakan oleh sistem yang ditujukan untuk pengguna. Berikut ini adalah fitur yang dapat digunakan di aplikasi ini :

Pengguna

1. Calon pengguna mendaftar ke dalam aplikasi
2. Pengguna bisa login
3. Pengguna bisa melihat data mentah batubara yang ada di dalam aplikasi dan memilih data yang ingin di cluster.
4. Pengguna bisa memasukkan jumlah cluster yang diinginkan
5. Pengguna bisa melihat hasil visualisasi cluster
6. Pengguna bisa melihat hasil analisa cluster

3.3 Perancangan Sistem

Perancangan sistem adalah perancangan susunan langkah-langkah agar suatu sistem bisa berjalan sesuai dengan fungsi dan alasan sistem itu dibuat. Dalam sub-bab ini, perancangan aplikasi dilakukan dengan pemodelan Data Flow Diagram (DFD) dan Entity Relation Diagram (ERD),

3.3.1 Diagram konteks

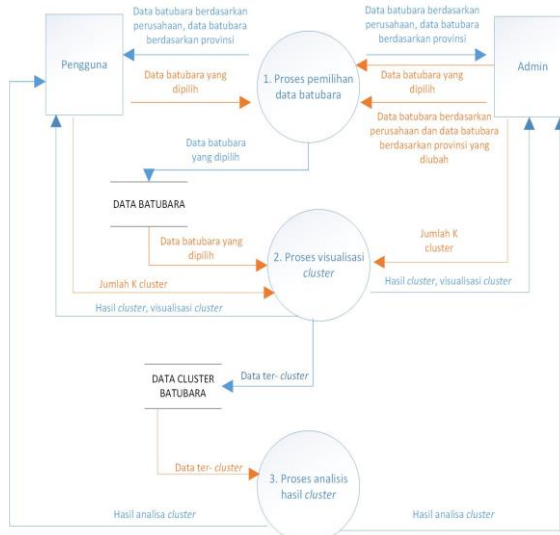
Diagram ini hanya menunjukkan satu proses yang merupakan proses sistem secara keseluruhan.



Gambar 3.1 Diagram konteks

3.3.2 Data Flow Diagram (DFD)

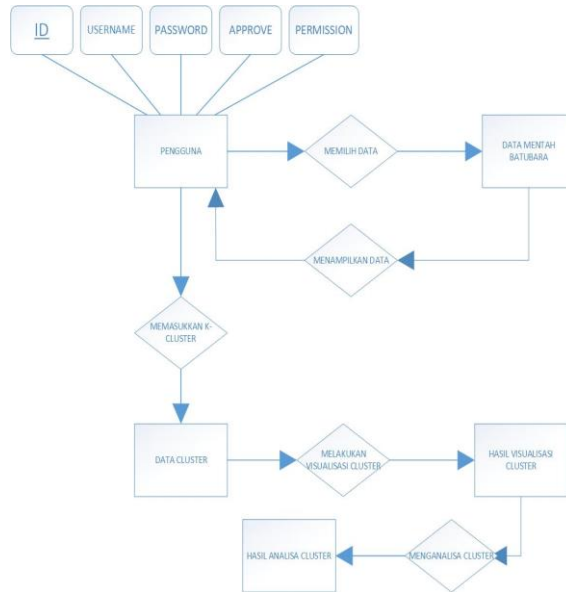
Data Flow Diagram digunakan sebagai pemodelan data untuk memberi gambaran asal dan tujuan data yang keluar dari sistem serta interaksi antara data tersebut.



Gambar 3.2 DFD level 1

3.3.3 Entity Relationship Diagram (ERD)

ERD adalah permodelan yang digunakan untuk menjelaskan relasi antar entitas dalam suatu sistem.

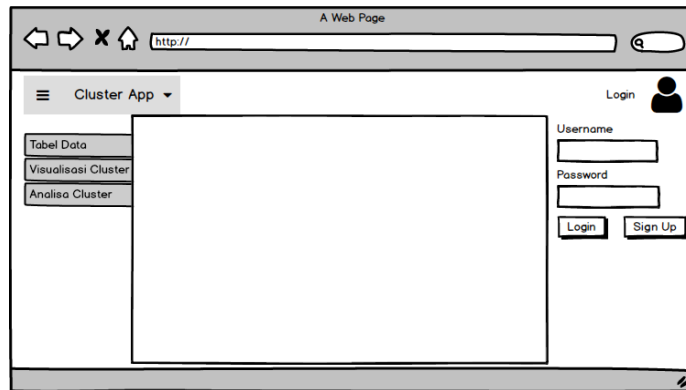


Gambar 3.3 ERD

Pada Gambar 3.6 ERD, terlihat bahwa pengguna memiliki 5 atribut yaitu : ID sebagai *key* atribut, *username*, *password*, *permission*, dan *approve*.

3.4 Desain Antarmuka

Desain antarmuka adalah desain dari suatu perangkat keras atau lunak yang dirancang untuk mempermudah pengguna dalam menggunakan perangkat dengan efisien dan mudah. Desain antarmuka dibuat agar pengguna bisa menggunakan aplikasi clustering ini dengan mudah dan efisien.

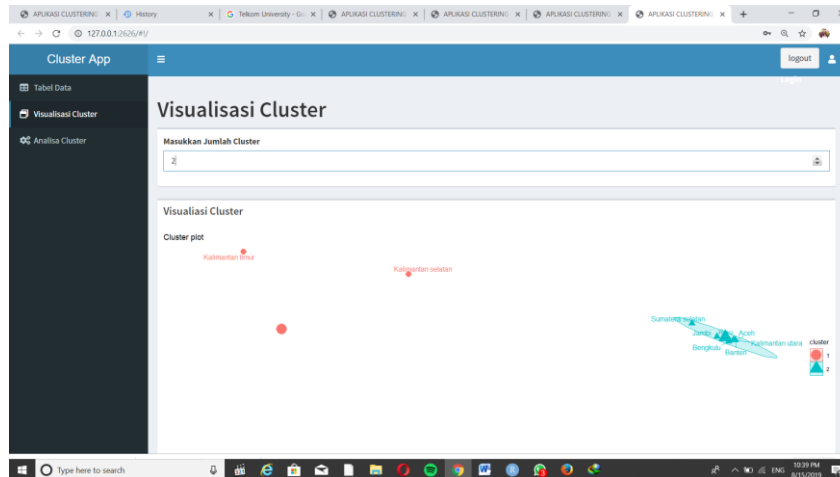


Gambar 3.4 Tampilan halaman awal

Gambar 3.7 halaman awal memperlihatkan tampilan awal ketika website dibuka. Untuk bisa mengakses aplikasi, pengguna diharuskan login terlebih dahulu

3.5 Implementasi

Aplikasi clustering berbasis web ini mengimplementasikan algoritma K-Means dalam pengelompokkan datanya dan hasil cluster dianalisa untuk melihat pola dari data. Jumlah nilai cluster yang bisa dimasukkan oleh pengguna adalah 2 hingga 10 cluster namun jumlah cluster yang digunakan sebagai k-cluster untuk dianalisa dibuku ini adalah 3 cluster.



Gambar 3.5 Menu visualisasi cluster

3.6 Pengujian Alpha

Pengujian alpha digunakan untuk mengevaluasi proses yang ada di dalam sistem apakah sudah sesuai dengan yang diinginkan oleh pengembang aplikasi atau belum. Pengujian alpha yang digunakan dapat dilihat di tabel 4.1 pengujian alpha

Tabel 3.1 Pengujian alpha

No	Fungsi yang diuji	Cara pengujian	Hasil yang diharapkan	Hasil pengujian
1.	Tabel data	Memilih data mentah berdasarkan jenis data dan tahunnya	Tampil hasil data mentah berdasarkan jenis dan tahun yang dipilih	Berhasil
2.	Visualisasi cluster	Memasukkan jumlah cluster mulai dari rentang 2-10	Tampil hasil visualisasi berdasarkan jumlah cluster yang dimasukkan dan berentang 2-10	Berhasil
3.	Analisa kluster	Menganalisa hasil cluster sesuai dengan jumlah cluster yang dimasukkan	Tampil analisa sesuai dengan jumlah cluster yang dimasukkan	Berhasil

3.7 Pengujian Beta

Hasil pengujian beta yang dilakukan terhadap 44 responden pada tanggal 4 Agustus 2019 dengan jumlah 8 pertanyaan, didapat hasil 33,3% sangat bermanfaat, 22% bermanfaat, 15,6% cukup bermanfaat, 2,2% kurang bermanfaat, dan 0% yang menilai sangat tidak bermanfaat

3.8 Analisa Data Hasil Clustering

3.8.1 Data Awal

Data yang digunakan berasal dari kementerian ESDM yang dipublikasikan secara online di website resmi mereka. Data yang dikelola di aplikasi ini ada dua yaitu data batubara berdasarkan perusahaan dan data batubara berdasarkan provinsi. Data batubara berdasarkan perusahaan memiliki 4 parameter yaitu produksi, ekspor, penjualan domestik, dan jumlah Karyawan. Data batubara berdasarkan provinsi juga memiliki 4 parameter yaitu produksi, ekspor, penjualan domestik, dan lisensi bisnis pertambangan. Data yang ditunjukkan pada bab ini hanyalah data batubara berdasarkan provinsi pada tahun 2014. Beberapa data di dalam e-book tidak diketahui nilainya atau N/A dikarenakan kebijakan dan kemampuan pengumpulan data batubara pada pemerintah daerah setempat masih belum ada yang diterima oleh pemerintah pusat.

Tabel 3.2 Data batubara berdasarkan provinsi tahun 2014

Provinsi	Produksi(Ton)	Penjualan domestik(Ton)	Lisensi bisnis pertambangan(per-buah)	Ekspor(Ton)
Kalimantan selatan	160398760.8	27325911.74	7646	130628810.8
Kalimantan timur	246090964.6	31404538.54	3918	206933158.9
Kalimantan tengah	9731450.64	771489	0	14651046
Kalimantan utara	7832482.57	15045	15	0
Sumatera selatan	26450196.47	15939681.32	6933	9712872.5
Aceh	127172	127172	127	0
Bengkulu	4267375.15	0	0	4220946.73
Jambi	2576730.15	564228	0	1920832.26
Riau	621574.58	31936	32	491948.55
Banten	0	0	0	0

Pada tabel 3.2 data batubara berdasarkan provinsi, provinsi yang ada pada data ini ada 10 provinsi yaitu Kalimantan Selatan, Kalimantan Utara, Kalimantan Tengah, Kalimantan Timur, Sumatera Selatan, Aceh, Bengkulu, Jambi, Riau dan Banten pada tahun 2014.

3.8.2 Hasil Cluster dan Analisanya

Data mentah yang sudah dipilih oleh pengguna di cluster secara otomatis oleh sistem.

Tabel 3.3 Analisa cluster batubara berdasarkan provinsi 2014

Cluster	Provinsi	Produksi(Ton)	Penjualan domestik(Ton)	Lisensi bisnis pertambangan(perbuah)	Ekspor(Ton)
3	Kalimantan Selatan	160398760.8	27325911.74	7646	130628810.8
3	Kalimantan Timur	246090964.6	31404538.54	3918	206933158.9
2	Kalimantan Tengah	9731450.64	771489	0	14651046
2	Kalimantan Utara	7832482.57	15045	15	0
1	Sumatera Selatan	26450196.47	15939681.32	6933	9712872.5
2	Aceh	127172	127172	127	0
2	Bengkulu	4267375.15	0	0	4220946.73
2	Jambi	2576730.15	564228	0	1920832.26
2	Riau	621574.58	31936	32	491948.55
2	Banten	0	0	0	0

Pada tabel 3.3 analisa *cluster* batubara berdasarkan provinsi tahun 2014 menunjukkan hasil data yang sudah ter *cluster*. Provinsi yang tergolong dalam *cluster* 1 adalah Aceh, Bengkulu, Jambi, Riau dan Banten. Provinsi yang tergolong dalam *cluster* 2 adalah Kalimantan Utara dan Sumatera Selatan. Provinsi yang tergolong dalam *cluster* 3 adalah Kalimantan Selatan, Kalimantan Tengah dan Kalimantan Timur.

Tabel 3.4 Rata-rata cluster

Cluster	Produksi(Ton)	Penjualan domestik(Ton)	Lisensi bisnis pertambangan(per-buah)	Ekspor(Ton)
1	26450196.47	15939681.32	6933	9712872.5
2	3593826.441	215695.7143	24.85714286	3040681.934
3	203244862.7	29365225.14	5782	168780984.8

Tabel 3.5 Maksimum cluster

Cluster	Produksi(Ton)	Penjualan domestik(Ton)	Lisensi bisnis pertambangan(per-buah)	Ekspor(Ton)
1	26450196.47	15939681.32	6933	9712872.5
2	9731450.64	771489	127	14651046
3	246090964.6	31404538.54	7646	206933158.9

Tabel 3.6 Minimum cluster

Cluster	Produksi(Ton)	Penjualan_domestik(Ton)	Lisensi bisnis pertambangan(per-buah)	Ekspor(Ton)
1	26450196.47	15939681.32	6933	9712872.5
2	0	0	0	0
3	160398760.8	27325911.47	3918	130628810.8

4. Kesimpulan

Berdasarkan tujuan dari Tugas Akhir Pengelompokan Data Batubara Di Indonesia Menggunakan Metode K-Means Clustering maka dapat ditarik kesimpulan sebagai berikut:

1. Telah dibuat Aplikasi berbasis web yang bisa mengelompokkan data batubara dengan metode K-Means Clustering.
2. Aplikasi ini bisa menampilkan visualisasi hasil K-Means Clustering.
3. Aplikasi ini bisa menampilkan hasil analisa K-Means Clustering.
4. Aplikasi ini mampu membantu kementerian ESDM dalam pengelompokkan data secara otomatis.

Daftar Pustaka:

- [1] Balasubramanian, A. 2017. *Coal Mining Methods*. Mysore : Universitas Mysore
- [2] The British Petroleum Company. 2017, *Statistical Review of World Energy Underpinning Data*. [available] <http://www.bp.com/statisticalreview>
- [3] Narwati, **Pengelompokkan Mahasiswa menggunakan Algoritma k-means**
- [4] Rahman, Tegar, Aulia, Wiranto dan Rini Anggraini.2017. *Coal Trade Clustering using K-Means(Case Study PT Global Bangkit Utama)*.
- [5] Kusuma, Daru, Purba. 2018. *Powerplant Clustering In Indonesia Using K-means*. Vol. X No. X, september 201X
- [6] J. Han, K, Micheline dan
- [7] J. Pei. 2012. *Data Mining : Concepts and Techniques (Third Edition)*, Waltham, MA: Morgan Kaufmann Publishers.
- [8] S. Russell and P. Norvig, *Artificial Intelligence A Modern Approach*. Upper Saddle River, New Jersey 07458: Pearson Education, Inc., 3 ed., 2010.
- [9] Kementerian energi dan sumber daya mineral. 2015. *Informasi mineral dan batubara*. Indonesia

