

Klasterisasi Data Penduduk Untuk Penerimaan Bantuan Pangan Non Tunai (BPNT) Menggunakan K-Means (Studi Kasus : Desa Tanimulya Bandung Barat)

Djuniardi Suhardinata¹, Ade Kania Ningsih², Fatan Kasyidi³

{djuniardis18@if.unjani.ac.id¹}

Universitas Jenderal Achmad Yani^{1,2,3}

Abstrak.

Dalam perkembangan yang terjadi di masyarakat desa tanimulya yang pesat terdapat pula pembagian beberapa kelompok masyarakat yang di golongkan dalam hal perekonomiannya untuk golongan yang masuk dalam kalangan menengah keatas dan menengah mungkin dapat bertahan dan terus berkembang, sedangkan untuk yang masuk dalam golongan kebawah, memerlukan bantuan dari orang lain, bantuan yang di butuhkan sudah di berikan, namun pembagian dari bantuan pangan non tunai (BPNT) ini, pemberiannya belum merata dan efektif, mengapa terjadi seperti itu karena data penduduk yang digunakan belum cukup mendukung untuk pembagian yang merata dan masih manual, oleh karena itu diperlukan sebuah sistem pengklusteran yang dimana nanti disesuaikan dengan aturan atau kategori yang sudah ada untuk mengelompokkan masyarakat agar nantinya penyalurannya bantuan dapat lebih maksimal lagi, klusterisasi yang merupakan sebuah proses untuk mengelompokkan data kedalam beberapa kluster dimanadi data dalam kluster memiliki kemiripan dengan sesama nya, data penduduk yang sudah di olah nantinya diharapkan dapat membantu dalam program bantuan pangan non tunai (BPNT) yang dijalankan saat ini.

Kata kunci – Data Penduduk, Klasterisasi, Klaster, Bantuan Pangan Non Tunai (BPNT)

1. Pendahuluan

Program bantuan sosial yang diberikan pemerintah telah pemerintah jalankan saat ini diantaranya dari Program Indonesia Pintar (PIP), Program Jaminan Kesehatan Nasional (JKN-KIS), Program Keluarga Harapan (PKH) dan Bantuan Pangan Non Tunai (BPNT), tujuan dari program bantuan tersebut yaitu diantaranya untuk menanggulangi masalah kemiskinan yang ada di Indonesia, menurut sumber kominfo.go.id bahwa angka langkah-langkah yang dilakukan pemerintah telah menurunkan angka kesejahteraan dari 11,22% pada tahun 2015 dan di tahun 2018 menjadi 9,82%. Program Indonesia Pintar untuk bantuan pendidikan anak-anak yang tidak mampu dari SD, SMP, SMA juga program bantuan Jaminan Kesehatan Nasional yang fokus pada kesehatan

masyarakat tidak mampu, Program Keluarga Harapan yang bertujuan untuk meningkatkan kesejahteraan dan tidak lupa program Bantuan Pangan Non Tunai dimana sasaran utama disini kepada keluarga yang terkena dampak dari tragedi COVID-19 dan disini saya memfokus pada program Bantuan Pangan Non Tunai dimana bantuan tersebut berjalan pada lingkungan saya sendiri.

Saat ini pada Desa Tanimulya sedang diadakannya program Bantuan Pangan Non Tunai (BPNT) yang penyalurannya masih kurang tepat, masih ditemukan masyarakat penerima bantuan yang masuk dalam ekonomi sangat sejahtera, dan ada juga masyarakat yang berekonomi tidak sejahtera tapi tidak mendapatkan bantuan, penyaluran bantuan yang tidak tepat akan menimbulkan kecemburuan sosial [2], hal ini menjelaskan kurang tepatnya penyaluran bantuan yang sedang di jalankan. Pada suatu daerah yang memiliki perbedaan dalam status ekonomi yang beraneka ragam, diperlukan adanya pengelompokan untuk dapat mengetahui perbedaan atau kelas dari status ekonomi masyarakat tersebut, diharapkan pengelompokan tersebut dapat membantu dalam program yang ditujukan untuk penyaluran Bantuan Pangan

Non Tunai (BPNT) yang di peruntukkan bagi masyarakat golongan menengah sampai kurang sejahtera.

Bantuan Pangan Non Tunai (BPNT) ini adalah bantuan yang disalurkan melalui Himpunan Bank Milik Negara (Himbara) atau melalui elektronik warong (e-Warong) sebesar Rp.200.000,- dana ini di pergunakan untuk memberi beras sebanyak 10Kg yang akan di berikan oleh gerai Bulog yang tersebar di seluruh wilayah Indonesia. Di sini, desa berperan sebagai penyalur bantuan kepada masyarakat di wilayahnya masing-masing. Adanya Bantuan Pangan Non-Moneter (BPNT) pemerintah yang penyalurannya masih menggunakan data basi tanpa memperbarui data yang sudah ada, yang kategorisasinya hanya mengikuti pedoman Dinas Sosial, artinya penyaluran pangan Bantuan Pangan Non Tunai (BPNT) akan tidak merata dan tidak akurat karena beberapa keluarga tidak menerima bantuan padahal mereka merupakan bagian dari keluarga penerima manfaat atau menerima bantuan dalam penyaluran berikutnya namun tidak diakui sehingga tidak menerimanya. Oleh karena itu, penyaluran atau pemberian bantuan pangan nontunai (NBNT) kurang tepat dan merata.

Penyelesaian permasalahan tersebut memerlukan informasi ilmiah yang membantu penyaluran dukungan yang baik, salah satu caranya adalah dengan menggunakan clustering. Clustering adalah suatu metode data mining yang secara otomatis mengelompokkan data tanpa pemberitahuan label kelas[4], kemudian mencari kesamaan pada data tersebut, jika ada kesamaan pada data penerima, hal ini menjadi salah satu faktor penentunya. . Pada penelitian ini metode clustering yang digunakan adalah metode k-means. Metode k-means merupakan algoritma clustering yang paling banyak digunakan dalam aplikasi karena kemudahannya [3], dalam mencari algoritma Naïve Bayes untuk mengklasifikasikan penerima bantuan pangan non-moneter menurut hasil Castaka Agus Sugianto [5] Berdasarkan pengujian dan validasi global algoritma Naïve Bayes, mencapai akurasi 58,29%, akurasi 92,90%, recovery 21,84%, AUC 0,765 dan ukuran F 34,42%, sedangkan algoritma pohon keputusan mencapai akurasi 73,97%, akurasi 85,04%. Recall 61,92%, AUC 0,746, F terukur 71,17%. Dari hasil tersebut dapat dikatakan bahwa algoritma pohon keputusan memiliki akurasi yang lebih tinggi dalam memprediksi data penerima manfaat dan

dalam penelitian taksonomi untuk mengidentifikasi riwayat penerima bantuan pangan non-moneter. Menggunakan metode k-means clustering Sulton Ghousi Pratama [9] yang dimaksudkan untuk digunakan data terkini ketika memeriksa apakah data yang digunakan dalam penelitian ini menggunakan data lama.

2. Metode

2.1 Data Mining

Data Mining diartikan sebagai menambang data atau upaya untuk menggali informasi yang berguna pada database yang sangat besar. Hal terpenting dalam teknik data mining adalah aturan untuk menemukan pola frekuensi tinggi antar himpunan item set yang disebut fungsi Association Rules (Aturan Asosiasi)[11]. Beberapa algoritma yang termasuk dalam Aturan Asosiasi adalah seperti AIS Algorithm, Apriori Algorithm, DHP Algorithm, dan Partition Algorithm [10].

2.2 Normalisasi Min-Max

Proses Data mining akan dapat menghasilkan hasil yang efektif dan tepat jika nilai yang dihasilkan memiliki keakuratan yang tinggi, salah satu teknik yang digunakan yaitu normalisasi, normalisasi digunakan untuk membakukan semua fitur dari dataset ke dalam kriteria yang telah ditentukan sebelumnya[14]

$v'_i = \left(\frac{v_i - \min_A}{\max_A - \min_a} * (new_max_A - new_min_A) \right) + new_min_A$	(1)
---	-------

2.3 Klasterrisasi

Klasterisasi merupakan salah satu teknik atau metode pada data mining. Algoritma klasterisasi merupakan algoritma pengelompokkan sejumlah data menjadi kelompok-kelompok data tertentu (klaster)[6]. Clustering adalah pengelompokan menggunakan teknik unsupervised learning dimana tidak diperlukan fase learning serta tidak menggunakan pelabelan pada setiap kelompok [6].

2.4 K-Means

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, unsupervised learing juga dimana hanya menginputkan data (x) dan tidak ada keluaran dari variable yang sesuai, dan tujuan dari *unsupervised learing* adalah untuk memodelkan struktur atau distribusi yang mendasari data untuk di pelajari lebih lanjut. Metode k-means ini mengelompokkan data yang ada kedalam beberapa kelompok yang masing-masing memiliki karakteristik[12].

Tujuan dari algoritma adalah menemukan jarak antara objek dengan *centroid* yang terdekat, yaitu dengan meminimumkan fungsi objektif J yang dirumuskan sebagai fungsi dari U dan V sebagai berikut :

$$J(U, V) = \sum_{k=1}^n \sum_{i=1}^c \mu_{ik} d^2(x_k, v_i)$$

2.5 Silhouette Coefficient

Evaluasi *Silhouette Coefficient* ini menggabungkan 2 metode, yaitu metode *Cohesion* dan metode *Separation*. Metode *Cohesion* digunakan untuk mengukur jarak antara objek satu dengan objek lainnya didalam sebuah *cluster*, sedangkan metode *Separation* digunakan untuk mengukur jarak dari *cluster* pertama hingga *cluster* lainnya [20].

Setelah semua nilai diketahui maka nilai *Silhouette Coefficient* dapat ditentukan dengan menggunakan persamaan berikut:

$S_i = \frac{b_i - a_i}{\max(a_i - b_i)}$	(2)
---	-------

3. Hasil dan pembahasan

3.1 Nilai Silhouette Coefficient

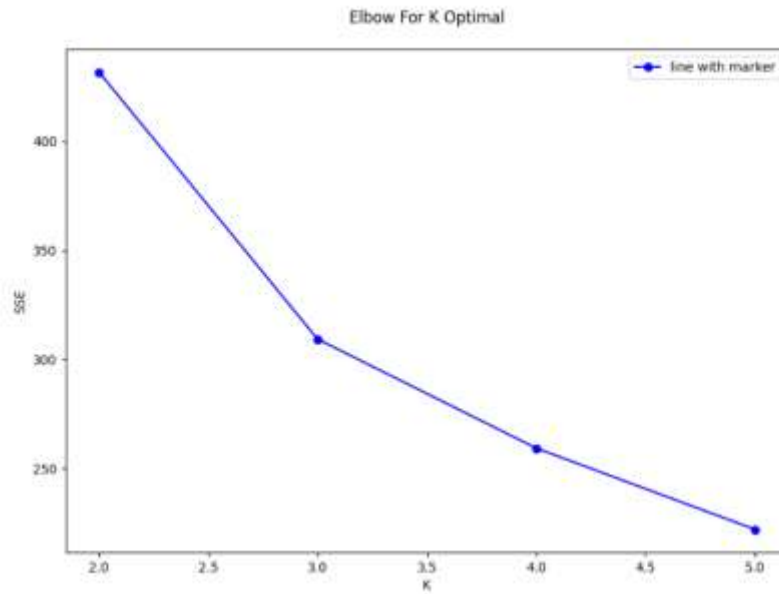
Perhitungan yang telah dilakukan dengan menggunakan metode K-Means selanjutnya dilakukan uji performance yang digunakan untuk melihat baik atau tidak nya hasil cluster, dilihat berdasarkan jarak antar obyek satu dengan objek lainnya didalam sebuah cluster dan jarak antar cluster dengan cluster lainnya. Hasil nilai *Silhouette Coefficient* didapat.

Show 10 entries

Nilai K	Silhouette Coefficient
3	0.44010439418634395
2	0.397446301505382

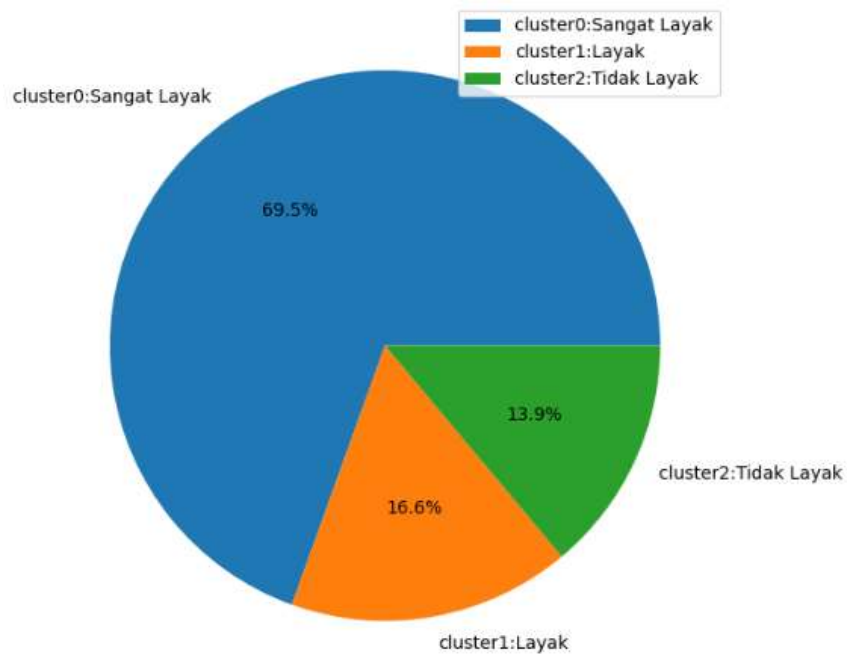
Gambar 1 Nilai Silhouette Coefficient

Berdasarkan hasil perhitungan yang telah dilakukan seperti pada Gambar 4 10, dapat dilihat bahwa nilai *Silhouette* pada cluster 3 yaitu 0.440104 Dimana dengan nilai yang didapat tersebut tergolong kedalam *cluster Weak Stucture*. Pada penelitian ini nilai K yang diambil yaitu 3 karena pada perhitungan elbow nilai yang memiliki sudut terkecil berada pada nilai K=3, oleh karena itu nilai K terbaik adalah K=3, untuk memperbesar nilai dari *silhouette* itu dibutuhkan attribute yang lebih banyak dan juga jumlah data yang lebih besar, agar dapat masuk kedalam *cluster Strong Structure*. Berikut grafik elbow yang dihasilkan pada penelitian ini dapat dilihat pada gambar.



Gambar 2 Hasil Elbow

Dari hasil cluster yang terbentuk yaitu 3 cluster yang berkelompok berdasarkan kemiripan dari polanya disini pengaruh tersebut di lihat dari value dari attribute data dan untuk hasil dapat di lihat pada gambar dibawah.



4. Kesimpulan

bedasarkan pada hasil penelitian yang dilakukan ini. Dihasilkannya beberapa *cluster*, dimana *cluster* optimal yang didapat menggunakan metode *elbow* adalah 3 dimana masing masing *cluster* menggambarkan aspek kemiripan antara data penerima bantuan dengan 3 kelompok, hasil tersebut *cluster 0* itu merupakan kumpulan penerima yang bisa dilihat kebanyakan itu data yang tinggi di attribute penghasilan, dan *cluster 1* merupakan kumpulan penerima yang bisa dilihat kebanyakan itu data yang tidak terlalu tinggi di attribute penghasilan tetapi di beberapa attribute itu nilainya sedang dan *cluster 2* merupakan kumpulan penerima yang bisa dilihat kebanyakan itu data yang tidak terlalu tinggi di attribute penghasilan tetapi di beberapa attribute itu nilainya kecil/rendah. Hasil nilai dari evaluasi dengan menggunakan metode *Silhouette Coefficient* untuk 3 *cluster* yaitu 0.440104394 sehingga hasil *cluster* yang terbentuk dapat dimasukkan kedalam kategori *cluster weak structure* dengan $range\ 0.25 < SC \leq 0.5$ dan merupakan *cluster* yang baik pada tingkat itu, Dari hasil evaluasi tersebut dapat di Tarik sebuah kesimpulan bahwa algoritma *K-Means* sangat baik dalam mengelompokan penerima bantuan, juga di yang tergambar dari hasil penelitian ini bahwa kemiripan dari pola yang di temukan itu berdasarkan pada nilai attribute yang di miliki setiap datanya.

Daftar Pustaka

- [1] ASRONI, HIDAYATUL FITRI, EKO PRASETYO. "Penerapan Metode Clustering dengan Algoritma K-Means pada Pengelompokkan Data Calon Mahasiswa Baru di Universitas Muhammadiyah Yogyakarta (Studi Kasus: Fakultas Kedokteran dan Ilmu Kesehatan, dan Fakultas Ilmu Sosial dan Ilmu Politik)." *SEMESTA TEKNIKA*, 2018.
- [2] Castaka Agus Sugianto, Firdi Rizky Maulana. "Algoritma Naïve Bayes Untuk Klasifikasi Penerima Bantuan Pangan Non Tunai (Studi KasusKelurahan Utama)." *Techno.COM*, 2019.
- [3] Fatmawati, Kiki. N.D. "Data Mining: Penerapan Rapidminer Dengan K-Means Klaster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi." *Cess (Journal Of Computer Engineering System And Science)*, n.d.
- [4] Haditsha. "Penerapan Data Mining Menentukan Strategi Penjualan Variasi Pakaian Menggunakan Metode K-Means Clustering." *JURNAL INFORMATIKA UPGRIS VOL 5*, 2019.
- [5] Helvine Gultom, Paulus Kindangen,George M.V. Kawung. "ANALISIS PENGARUH PROGRAM BANTUAN PANGAN NON TUNAI (BPNT) DAN PROGRAM KELUARGA HARAPAN (PKH) TERHADAP KEMISKINAN DI KABUPATEN MINAHASA TENGGARA." *Pembangunan Ekonomi dan Keuangan Daerah*, 2020.
- [6] Hendro Priyatman, Fahmi Sajid,Dannis Haldivany. "Klasterisasi Menggunakan

- Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan Mahasiswa." *Edukasi dan Penelitian Informatika*, 2019.
- [7] Rofiqo, Nurul. "Penerapan Clustering Pada Penduduk Yang Mempunyai Keluhan Kesehatan Dengan Datamining K-Means." *Komik (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 2018.
- [8] Saut Parsaoran Tamba, Felix Toknady Kesuma, Feryanto. "Penerapan Data Mining Untuk Menentukan Penjualan Sparepart Toyota Dengan Metode K-Means Clustering." *Jurnal Sistem Informasi Ilmu Komputer Prima (Jusikom Prima)*, 2019: 67-72.
- [9] Sulton Ghousi Pratama, Ali Mahmudi, Sentot Achmadi. "KLASIFIKASI PENENTUAN PENERIMA BANTUAN PANGAN NON TUNAI MENGGUNAKAN METODE K-MEANS CLUSTERING." *JATI (Jurnal Mahasiswa Teknik Informatika)*, 2020.
- [10] Tampubolon, Kennedy. "Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan." *Informasi Dan Teknologi Ilmiah (Inti)*, 2016.
- [11] Windarto, Agus Perdana. "Penerapan Data Mining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means." *Techno.Com*, 2017: 348-357 348-249.
- [12] B. M. Metisen And H. L. Sari, "Analisis Clustering Menggunakan Metode K-Means Dalam Pengelompokkan Penjualan Produk Pada Swalayan Fadhila," *J. Media Infotama*, Vol. 11, No. 2, Pp. 110–118, 2015
- [13] W. A. N. G. Zhao-Yi, H. U. A. N. G. Zheng-De, Y. A. N. G. Ping, R. E. N. Ting, and L. I. Xin-Hui, "Regularity of Wind-dispelling Medication Prescribed by LI Dong-Yuan: A Data Mining Technology-based Study," *Digit. Chinese Med.*, vol. 3, no. 1, pp. 20–33, 2020, doi: 10.1016/j.dcmmed.2020.03.003
- [14] Y. Wautelet, "Using the RUP/UML business use case model for service development governance: A business and IT alignment based approach," *Proc. - 2020 IEEE 22nd Conf. Bus. Informatics, CBI 2020*, vol. 2, pp. 121–130, 2020, doi: 10.1109/CBI49978.2020.10069.
- [15] C. Djaoui, E. Kerkouche, K. Khalfaoui, and A. Chaoui, "A graph transformation approach to generate analysable maude specifications from UML interaction overview diagrams," *Proc. - 2018 IEEE 19th Int. Conf. Inf. Reuse Integr. Data Sci. IRI 2018*, pp. 511–517, 2018, doi: 10.1109/IRI.2018.00081.
- [16] M. Bhuiyan, "Integration of Organisational Models and UML Use Case," *J. Comput.*, vol. 13, no. 1, pp. 1–17, 2018, doi: 10.17706/jcp.13.1.1-17.
- [17] T. A. Kurniawan, "Pemodelan Use Case (UML): Evaluasi Terhadap beberapa Kesalahan dalam Praktik," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 1, p. 77, 2018, doi: 10.25126/jtiik.201851610.
- [18] M. Elsayed, N. Elkashef, and Y. F. Hassan, "Mapping UML sequence diagram into the web ontology language OWL," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 5, pp. 318–326, 2020, doi: 10.14569/IJACSA.2020.0110542.
- [19] M. H. L. Vo and Q. Hoang, "Transformation of uml class diagram into owl ontology," *J. Inf. Telecommun.*, vol. 4, no. 1, pp. 1–16, 2020, doi:

10.1080/24751839.2019.1686681.

- [20] R. Nainggolan, R. Perangin-Angin, E. Simarmata, and A. F. Tarigan, "Improved the Performance of the K-Means Cluster Using the Sum of Squared Error (SSE) optimized by using the Elbow Method," *J. Phys. Conf. Ser.*, vol. 1361, no. 1, 2019, doi: 10.1088/1742-6596/1361/1/012015.
- [21] T. K. Barik and V. A. Centeno, "K-Medoids Clustering of Setting Groups in Directional Overcurrent Relays for Distribution System Protection," 2020 IEEE Kansas Power Energy Conf. KPEC 2020, 2020, doi: 10.1109/KPEC47870.2020.9167531.
- [22] J. Zhou, Y. Pan, C. L. P. Chen, D. Wang, and S. Han, "K-medoids method based on divergence for uncertain data clustering," 2016 IEEE Int. Conf. Syst. Man, Cybern. SMC 2016 - Conf. Proc., pp. 2671–2674, 2017, doi: 10.1109/SMC.2016.7844643.
- [23] H. B. Zhou and J. T. Gao, "Automatic method for determining cluster number based on silhouette coefficient," *Adv. Mater. Res.*, vol. 951, pp. 227–230, 2014, doi: 10.4028/www.scientific.net/AMR.951.227.
- [24] S. Luan, X. Kong, B. Wang, Y. Guo, and X. You, "Silhouette coefficient based approach on cell-phone classification for unknown source images," *IEEE Int. Conf. Commun.*, pp. 6744–6747, 2012, doi: 10.1109/ICC.2012.6364928