



Artikel Penelitian

Optimasi Pengelompokan Data Pada Metode *K-Means* dengan Analisis *Outlier*

Pasek Agus Ariawan^{a,*}^aProgram Studi Magister Teknik elektro, Fakultas Teknik, Universitas Udayana, Bali, Indonesia

INFORMASI ARTIKEL

Sejarah Artikel:

Diterima Redaksi: 15 Juli 2019

Revisi Akhir: 04 September 2019

Diterbitkan Online: 19 September 2019

KATA KUNCI

*Clusterig,**K-Means,*

Optimasi,

Outlier

KORESPONDENSI

E-mail: pasekagusariawan38@gmail.com

A B S T R A C T

Data mining secara umum adalah proses analisis dan eksplorasi sejumlah besar data yang berbeda untuk menemukan pola yang bermakna. Berbagai teknik tersedia dalam *data mining* untuk ekstraksi pengetahuan antara lain klasifikasi, prediksi, estimasi, asosiasi dan pengelompokan. Salah satu metode pengelompokan yang terkenal adalah metode *K-Means* karena kesederhanaan dan efisiensinya, sehingga metode ini mudah diaplikasikan di berbagai bidang. Teknik pra-pemrosesan data diterapkan pada data mentah untuk membuat data bersih, bebas *noise*, dan konsisten, dalam *data mining* terdapat beberapa metode yang dapat digunakan tergantung dari tujuan yang akan dicapai. Deteksi *outlier* adalah penelitian dari *data mining* untuk mendeteksi kasus – kasus abnormal pada *database*. penelitian ini menggunakan metode *local outlier factor* untuk deteksi *outlier* dengan tujuan untuk mengatasi kekurangan dari metode *K-Means* yang sangat dipengaruhi oleh *noise* dan meningkatkan kualitas hasil cluster dari metode *K-Means*. Hasil dari penelitian, penggunaan metode *local outlier factor* pada data capaian kinerja pegawai dapat mendeteksi data *outlier* sebesar 22.8%. Rata – rata iterasi, nilai SSE dan waktu komputasi untuk satu kali konvergen pada metode *K-Means* dengan analisis *outlier* memberikan hasil yang lebih baik untuk jumlah *cluster* 2 – 10 dibandingkan dengan metode *K-Means* tradisional.

1. PENDAHULUAN

Data dapat mengandung berbagai macam informasi yang dapat digunakan untuk berbagai keperluan seperti dalam pengambilan keputusan, interpretasi masa depan, perilaku pelanggan dan sebagainya. Dengan kemajuan teknologi saat ini, sejumlah data tidak hanya dapat dihasilkan, melainkan dapat juga dikomunikasikan dan disimpan melalui *website*. Selain *website*, sumber data lain yang tersedia adalah situs sosial media, halaman *web*, dll. Bidang yang berhubungan dengan ekstraksi pengetahuan dari data yang tersedia dikenal sebagai *data mining*. *Data mining* secara umum adalah proses analisis dan eksplorasi sejumlah besar data yang berbeda untuk menemukan pola yang bermakna. Berbagai teknik tersedia dalam *data mining* untuk ekstraksi pengetahuan antara lain klasifikasi, prediksi, estimasi, asosiasi dan pengelompokan. Dalam pengelompokan, *cluster* diatur sedemikian rupa sehingga objek dengan sifat yang serupa terletak pada satu cluster atau dengan kata lain objek dalam dua kelompok yang berbeda dapat memiliki sifat yang berbeda. Berbagai jenis pengelompokan adalah berbasis partisi, berbasis hierarkis, berbasis kepadatan, berbasis kisi dan berbasis model [1]. Salah satu metode pengelompokan yang terkenal adalah metode *K-Means* karena

kesederhanaan dan efisiensinya, sehingga metode ini mudah diaplikasikan di berbagai bidang [2], seperti penelitian yang dilakukan sebelumnya [3] tentang penerapan metode *K-Means* dalam peningkatan mutu pembelajaran pada perguruan tinggi. Algoritma *K-Means* efektif dalam menghasilkan *cluster* untuk banyak aplikasi praktis. Algoritma ini menghasilkan berbagai jenis *cluster* tergantung pada pilihan acak *centroid* awal. Teknik pra-pemrosesan data diterapkan pada data mentah untuk membuat data bersih, bebas *noise*, dan konsisten. Normalisasi Data mengubah data mentah dengan mengubahnya menjadi rentang spesifik menggunakan transformasi linear yang dapat menghasilkan suatu *cluster* dengan kualitas yang baik, serta meningkatkan akurasi dari algoritma pengelompokan tersebut. Terdapat beberapa metode *data mining* yang dapat digunakan tergantung dari tujuan yang ingin dicapai. Deteksi *outlier* sebenarnya menemukan *anomaly* adalah penelitian dari untuk mendeteksi kasus – kasus abnormal pada *database*[4]. Menurut[5] dalam penelitian yang dilakukannya, metode *K-Means* dapat memetakan hasil produktivitas panen padi berdasarkan karakteristik datanya, hasil dari penelitian dapat memetakan data ke dalam 3 kelompok yaitu kurang dari target, sesuai target, dan melebihi target. Penelitian yang dilakukan sebelumnya[6] tentang rekomendasi kelompok belajar siswa dengan menggunakan teknik *outlier* di SDN pagelaran II, pada penelitian

ini menggunakan metode *local outlier factor* untuk mendeteksi *outlier* pada data siswa kelas 6. Penelitian yang dilakukan oleh [7] ini mencakup studi perbandingan tentang berbagai algoritma deteksi *outlier* seperti deteksi *outlier* berbasis statistik, deteksi *outlier* berbasis kedalaman, teknik berbasis *Clustering*, deteksi *outlier* berbasis kepadatan dll. Studi perbandingan metode deteksi *outlier* ini dilakukan untuk mengetahui mana dari algoritma deteksi *outlier* yang lebih dapat diterapkan pada data dimensi tinggi. Penelitian yang dilakukan [8] bertujuan untuk mendeteksi pencilan pada set data campuran. Selanjutnya, mengkombinasikan algoritma dari beberapa teknik seperti *subset cluster*, *outlier detection cluster based*, dan *Multiattribute Decision Making (MADM)*. Sebuah tes dilakukan terhadap set data dari *UCI Machine Learning Repository*. Evaluasi dilakukan untuk membandingkan rata-rata jangkauan yang mencapai rasio tertinggi antara jumlah pencilan yang tepat dan jumlah data. Dari tes, dapat disimpulkan bahwa MixCBLOF efektif untuk dideteksi *outlier* di set campuran data. Penelitian yang dilakukan [9] menyajikan solusi terdistribusi pertama untuk metode *Local Outlier Factor (LOF)* - teknik deteksi *outlier* populer yang terbukti sangat efektif untuk *dataset* yang besar. Seiring bertambahnya ukuran *dataset* secara radikal, algoritma LOF yang sangat skalabel yang memanfaatkan infrastruktur terdistribusi modern diperlukan. Ini menimbulkan tantangan signifikan karena kompleksitas definisi LOF, dan kurangnya akses ke seluruh *dataset* pada setiap proses penghitungan. Solusi kami menampilkan kerangka kerja pipa LOF terdistribusi, yang disebut DLOF. Penelitian lain tentang deteksi *outlier* dilakukan oleh [10] yang menyebutkan bahwa tujuan utama dari tahap data *preprocessing* adalah untuk mendeteksi *outlier*. Penelitian [11] bertujuan untuk mengatasi kekurangan metode *K-Means* yang sangat dipengaruhi oleh *noise*, penelitian yang dilakukan dengan mengimprovisasi metode *K-Means* dengan metode deteksi *noise* percobaan dilakukan pada set data yang besar dengan hasil *clustering* ditingkatkan secara signifikan dan dampak data *noise* pada algoritma *K-Means* berkurang secara efektif dan hasil *clustering* lebih akurat.

Berdasarkan pemaparan penelitian [7], [10], [11], [14], [15] yang sudah dilakukan oleh peneliti terdahulu, maka pada penelitian ini menggunakan metode *local outlier factor* untuk deteksi *outlier* dengan tujuan untuk mengatasi kekurangan dari metode *K-Means* yang sangat dipengaruhi oleh *noise* dan meningkatkan kualitas hasil *cluster* dari metode *K-Means*.

2. METODE

2.1. Data Penelitian

Penelitian ini menggunakan data sekunder yang berasal dari laporan capaian kinerja pegawai selama tahun 2017 di Univeristas Udayana. Fitur yang digunakan adalah nomor induk pegawai, nilai perilaku, nilai kehadiran, nilai kinerja. Ekstraksi fitur dilakukan dengan cara mencari rata-rata nilai nilai perilaku, nilai kehadiran dan nilai kinerja.

Tabel 1. Potongan Data *Cluster*

No	NIP	Perilaku	Kehadiran	Kinerja
1	'195403071980031004	93.67	100	155.372
2	'195612311982111001	81.6936	74.4136	125.395
3	'195710131980031003	95.094	74.232	94.036
4	'195712311979011005	89.8282	87.0991	152.6205
5	'195809051981031003	87.895	90.9625	168
6	'195812311983032009	86.8811	95.8778	156.09
7	'195907271984031001	67.365	64.2213	84
8	'195909241986032003	87.855	75.335	168
9	'195910051981031004	90.895	62.1825	115.354
10	'195911251982112001	78.99	82.625	141.6575

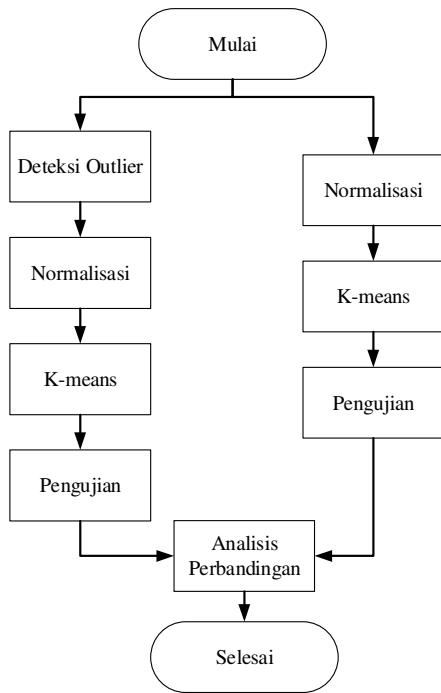
2.2. Tahapan Analisis

Penelitian ini menggunakan data yang berasal dari laporan capaian kinerja pegawai selama tahun 2017 di Univeristas Udayana. Fitur yang digunakan adalah nomor induk pegawai, nilai perilaku, nilai kehadiran, nilai kinerja.

Berdasarkan gambar 1 penelitian akan dilakukan dengan membandingkan metode *K-Means* menggunakan metode *local outlier factor* untuk proses dekteksi *outlier* dengan metode *K-Means* tradisional. Adapun tahapan normalisasi adalah untuk menyamakan rentang nilai antara fitur yang akan diuji.

Tahapan pengujian akan dilakukan dengan melihat jumlah iterasi, nilai SSE dan waktu proses konvergen pada metode *K-Means* menggunakan metode *local outlier factor* untuk proses dekteksi *outlier* dan metode *K-Means* tradisional. Pengujian dengan metode SSE dilakukan untuk pengukuran kualitas dari *cluster* yang terbentuk.

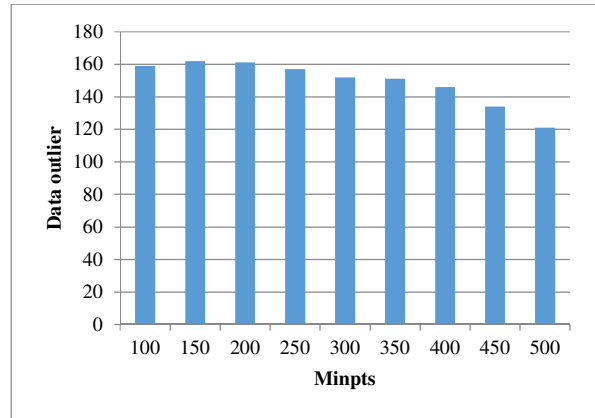
Tahapan analisis perbandingan akan dilakukan dengan membandingkan jumlah iterasi, nilai SSE dan waktu proses konvergen pada metode *K-Means* menggunakan metode *local outlier factor* untuk proses dekteksi *outlier* dan metode *K-Means* tradisional, kemudian menentukan pengaruh dari metode LOF pada proses klasterisasi. Pengujian dilakukan sebanyak 3 kali pengulangan, kemudian dicari rata-rata dari jumlah iterasi, nilai SSE dan waktu proses konvergen.



Gambar 1. Alur Proses Penelitian.

Pengujian dilakukan pada jumlah minpts yang telah ditentukan yaitu 100, 150, 200, 250, 300, 350, 400, 450, 500. Hasil deteksi outlier dapat dilihat pada tabel 2.

Dari tabel dapat dibuat grafik hasil deteksi outlier dengan metode local outlier factor. Gambar 2 menunjukkan hasil deteksi outlier dengan metode *local outlier factor*, dimana menunjukkan hasil deteksi tertinggi pada minpts 150 sebesar 23%, kemudian mengalami penurunan hasil deteksi hal ini diakibatkan oleh Semakin tingginya nilai minpts berarti semakin luas pula kepadatan objek tersebut. Data normal pada hasil deteksi dengan nilai minpts 150 kemudian akan digunakan untuk proses *clustering*.



Gambar 2. Grafik Hasil Deteksi Outlier

3. HASIL DAN PEMBAHASAN

3.1. Hasil Deteksi Outlier

Pada algoritma *local outlier factor* ini, hal-hal yang diperhatikan adalah menentukan banyaknya nilai minpts, minpts adalah jumlah tetangga terdekat yang akan digunakan untuk menghitung nilai outlier. Kemudian menentukan nilai batas dalam memilih suatu anomaly.

Tabel 2. Hasil Deteksi Outlier.

Minpts	Outlier		Normal	
	Data	Presentase	Data	Presentase
100	159	22.6%	546	77.4%
150	162	23.0%	543	77.0%
200	161	22.8%	544	77.2%
250	157	22.3%	548	77.7%
300	152	21.6%	553	78.4%
350	151	21.4%	554	78.6%
400	146	20.7%	559	79.3%
450	134	19.0%	571	81.0%
500	121	17.2%	584	82.8%

Pengujian outlier dilakukan untuk melakukan filter pada data – data yang memiliki karakteristik yang berbeda dari data lainnya.

3.2. Normalisasi Min-Max

Normalisasi bertujuan untuk menskalakan nilai dari fitur yang digunakan pada penelitian ini memiliki rentang yang sama dari 0-1. Tabel 3 merupakan potongan data awal sebelum dilakukan proses transformasi. Tabel 3 menunjukkan hasil dari transformasi yang telah dilakukan.

Dari tabel 3 kemudian dilakukan perhitungan dengan menggunakan rumus 3. Hasil dari perhitungan ditampilkan pada tabel 4, dimana dapat dilihat terjadi perubahan nilai pada setiap fitur yang digunakan pada penelitian, yang semula memiliki rentang nilai yang berbeda, kemudian setelah proses transformasi memiliki rentang nilai yang sama.

Tabel 3. Potongan Data Awal Sebelum Transformasi.

Perilaku	Kehadiran	Kinerja
93.67	100	155.372
81.69364	74.41364	125.395
95.094	74.232	94.036
89.82818	87.09909	152.6205
87.895	90.9625	168
86.88111	95.87778	156.09
67.365	64.22125	84
87.855	75.335	168

Tabel 4. Potongan Data Awal Sebelum Transformasi.

Perilaku	Kehadiran	Kinerja
0.939794	1	0.918498
0.757167	0.707985	0.725024
0.961509	0.705912	0.522631
0.88121	0.852763	0.90074
0.851731	0.896856	1
0.836271	0.952953	0.923132
0.53867	0.59166	0.457858
0.851121	0.7185	1

3.3. Hasil Pengujian Rata- Rata Jumlah Iterasi Dalam Satu Kali Konvergen Pada K-Means + LOF Dengan K-Means Tradisional

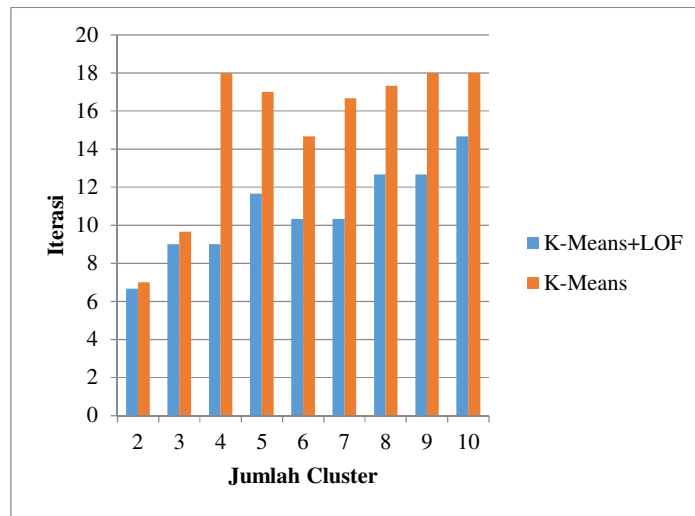
Pada pengujian pertama dilakukan untuk membandingkan rata – rata jumlah iterasi yang diperlukan untuk mencapai kondisi konvergen antara metode *K-Means* menggunakan metode *local outlier factor* untuk proses dekteksi *outlier* dengan *K-Means* tradisional.

Pengujian dilakukan sebanyak tiga kali percobaan dengan perubahan jumlah *cluster* yang diuji. Pengujian dilakukan pada jumlah *cluster* 2-10. Hasil rata – rata jumlah iterasi dapat dilihat pada tabel 5.

Tabel 5. Hasil Perbandingan Rata-Rata Iterasi

K	Iterasi	
	<i>K-Means</i> +LOF	<i>K-Means</i>
2	6.666666667	7
3	9	9.666666667
4	9	18
5	11.666666667	17
6	10.333333333	14.666666667
7	10.333333333	16.666666667
8	12.666666667	17.333333333
9	12.666666667	18
10	14.666666667	18

Dari tabel 5 dapat dibuat grafik hasil perbandingan uji coba rata-rata jumlah iterasi yang diperlukan untuk satu kali konvergen.



Gambar 3. Grafik Perbandingan Hasil Iterasi

Gambar 3 menunjukkan hasil metode *K-Means* dengan analisis *outlier* memiliki rata – rata nilai iterasi yang lebih kecil dari metode *K-Means* tradisional. Hal ini menunjukkan bahwa jumlah pengulangan pemetukan *cluster* untuk mencapai kondisi konvergen dengan metode *K-Means* dengan analisis *outlier* lebih cepat dibandingkan metode *K-Means* tradisional.

3.4. Hasil Pengujian Rata- Rata Nilai Sum Of Square Error Dalam Satu Kali Konvergen Pada K-Means + LOF Dengan K-Means Tradisional

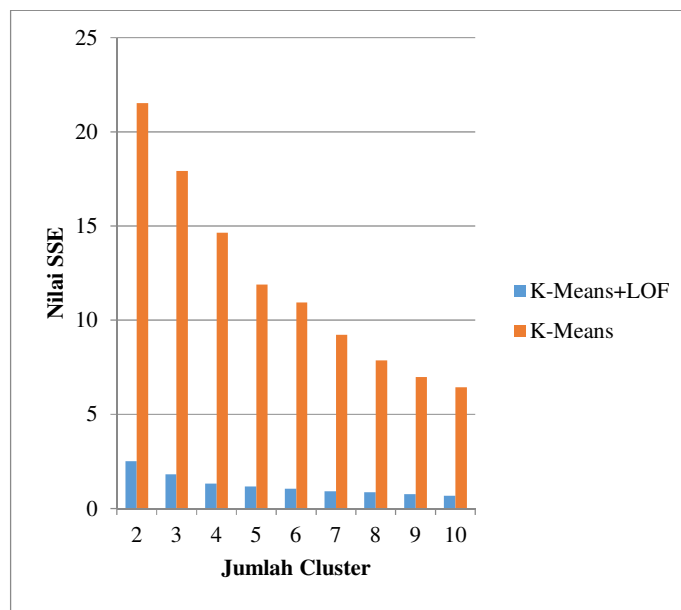
Pada pengujian kedua dilakukan untuk membandingkan rata – rata hasil *Sum of Square error* pada kondisi satu kali konvergen antara metode *K-Means* menggunakan metode *local outlier factor* untuk proses dekteksi *outlier* dengan *K-Means* tradisional. Pengujian dilakukan sebanyak tiga kali percobaan dengan perubahan jumlah

cluster yang diuji. Pengujian dilakukan pada jumlah cluster 2-10. Hasil rata – rata hasil *Sum of Square error* dapat dilihat pada tabel 6.

Dari tabel 6 dapat dibuat grafik hasil perbandingan uji coba rata – rata nilai *Sum of Square error* pada kondisi satu kali konvergen. Gambar 4 menunjukkan hasil *K-Means* dengan analisis *outlier* memiliki rata – rata nilai *Sum of Square error* yang lebih kecil dibandingkan *K-Means* tradisional. Ini menunjukkan bahwa entitas yang masuk pada jumlah cluster metode *K-Means* dengan analisis *outlier* memiliki kesamaan yang lebih seragam dari pada entitas yang ada pada jumlah cluster *K-Means* tradisional.

Tabel 6. Hasil Perbandingan *Sum Of Square Error*

K	<i>Sum of Square error</i>	
	<i>K-Means+LOF</i>	<i>K-Means</i>
2	2.500266667	21.5294
3	1.824433333	17.916867
4	1.329533333	14.647333
5	1.1644	11.886333
6	1.0572	10.940833
7	0.915866667	9.2099667
8	0.856533333	7.8622667
9	0.766033333	6.972
10	0.676366667	6.4364333



Gambar 4. Grafik Perbandingan Hasil *Sum Of Square Error*

3.5. Hasil Pengujian Rata- Rata Waktu Dalam Satu Kali Konvergen Pada K-Means + LOF Dengan K-Means Tradisional

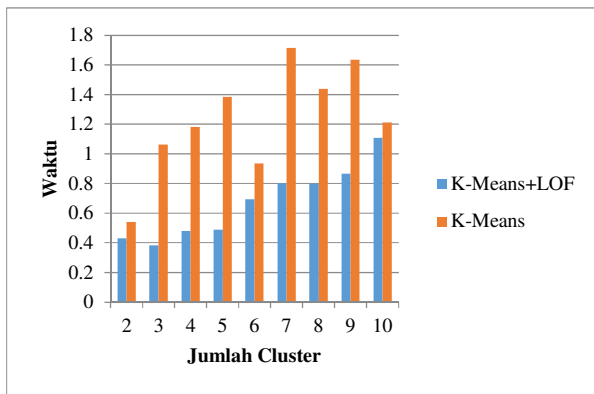
Pada pengujian Ketiga dilakukan untuk membandingkan rata – rata waktu yang diperlukan untuk kondisi satu kali konvergen antara metode *K-Means* menggunakan metode *local outlier factor* untuk proses dekteksi *outlier* dengan *K-Means* tradisional. Pengujian dilakukan sebanyak tiga kali percobaan dengan perubahan jumlah cluster yang diuji. Pengujian dilakukan pada jumlah cluster 2-10. Hasil rata – rata waktu dapat dilihat pada tabel 7.

Tabel 7. Hasil Perbandingan Waktu

K	Waktu	
	<i>K-Means+LOF</i>	<i>K-Means</i>
2	0.431266333	0.540047333
3	0.382809667	1.061490333
4	0.480251333	1.182051667
5	0.488963667	1.385132667
6	0.693024333	0.933994
7	0.800864333	1.713181667
8	0.798506333	1.438461667
9	0.864796333	1.635548
10	1.106676667	1.213016667

Dari tabel 7 dapat dibuat grafik hasil perbandingan uji coba rata – rata waktu yang diperlukan untuk kondisi satu kali konvergen.

Gambar 5 menunjukkan hasil *K-Means* dengan analisis *outlier* memiliki rata – rata waktu untuk mencapai satu kali konvergen yang lebih kecil dibandingkan *K-Means* tradisional. Ini menunjukkan bahwa waktu komputasi yang diperlukan metode *K-Means* dengan analisis *outlier* lebih cepat daripada *K-Means* tradisional dikarenakan data yang memiliki karakteristik yang jauh berbeda telah dihilangkan pada proses pengelompokan data.



Gambar 5. Grafik Perbandingan Waktu Untuk Satu Kali Konvergen.

4. KESIMPULAN

Berdasarkan hasil dari penelitian dapat disimpulkan bahwa Penggunaan metode *local outlier factor* pada data capaian kinerja pegawai dapat mendeteksi data *outlier* sebesar 22.8% dan dengan menggunakan metode *local outlier factor* dapat memberikan hasil yang lebih optimal pada pengelompokan data dengan metode *K-Means* baik dari jumlah iterasi, waktu pemrosesan dan kualitas *cluster*.

Metode *local outlier factor* untuk dekteksi outlier sangat sensitifiif terhadap nilai parameter minpts untuk menentukan jumlah tetangga terdekat. Oleh karena itu diperlukan suatu metode yang dapat menentukan jumlah minpts yang optimal pada metode *local outlier factor*.

DAFTAR PUSTAKA

[1] V. Bhatt, M. Dhakar, Dan B. K. Chaurasia, “Filtered Clustering Based On Local Outlier Factor In Data Mining,” *Int. J. Database Theory Appl.*, Vol. 9, No. 5, Hal. 275–282, 2016.

[2] G. Gustientiedina, M. H. Adiya, Dan Y. Desnelita, “Jurnal Nasional Teknologi Dan Sistem Informasi

Penerapan Algoritma *K-Means* Untuk Clustering Data Obat-Obatan Pada Rsud Pekanbaru,” *J. Nas. Teknol. Dan Sist. Inf.*, Vol. 05, No. 01, Hal. 17–24, 2019.

[3] K. Handoko, “Penerapan Data Mining Dalam Meningkatkan Mutu Pembelajaran Pada Instansi Perguruan Tinggi Menggunakan Metode K - Means Clustering (Studi Kasus Di Program Studi Tkj Akademi Komunitas Solok Selatan),” Vol. 02, No. 03, Hal. 31–40, 2016.

[4] K. G. Sharma, Y. Singh, Dan A. K. Srivastava, “Variance On Outlier Factor,” In *Impact*, 2017, Hal. 101–103.

[5] M. R. Ridlo, S. Defiyanti, Dan A. Primajaya, “Implementasi Algoritme *K-Means* Untuk Pemetaan Produktivitas Panen Padi Di Kabupaten Karawang,” In *Citee 2017*, 2017, Hal. 426–433.

[6] N. Idham, “Penerapan Outlier Analysis Sebagai Salah Satu Rekomendasi Kelompok Belajar Terhadap Siswa Kelas 6 Di Sdn Pagelaran Ii Program Studi Teknik Informatika,” Universitas Komputer Indonesia, 2017.

[7] K. Kaur Dan A. Garg, “Comparative Study Of Outlier Detection Algorithms,” *Int. J. Comput. Appl.* (0975, Vol. 147, No. 9, Hal. 21–26, 2016.

[8] D. Maryono Dan A. Djunaidy, “Berbasis Klaster Pada Set Data Dengan Atribut Campuran Numerik Dan Kategorikal *,” *J. Ilm. Kursor*, Vol. 5, No. 3, Hal. 197–204, 2010.

[9] Y. Yan, L. Cao, C. Kuhlman, Dan E. Rundensteiner, “Distributed Local Outlier Detection In Big Data,” In *Kdd 2017 Research Paper*, 2017, Hal. 1225–1234.

[10] A. Mahendra, “Pentapisan Dan Deteksi Data Outlier Dalam Proses Sistem Akusisi Data Pada Proses Sintering,” *Arsitron*, Vol. 6, No. 1, Hal. 1–7, 2015.

[11] J. Wang Dan X. Su, “An Improved *K-Means* Clustering Algorithm,” In *2011 Ieee 3rd International Conference On Communication Software And Networks*, 2011, Hal. 44–46.

[12] G. Ngurah, W. Paramartha, D. E. Ratnawati, Dan A. W. Widodo, “Analisis Perbandingan Metode *K-Means* Dengan Improved Semi- Supervised *K-Means* Pada Data Indeks Pembangunan Manusia (Ipm),” Vol. 1, No. 9, Hal. 813–824, 2017.

[13] R. Pamula, J. K. Deka, Dan S. Nandi, “An Outlier Detection Method Based On Clustering,” In *Second*

International Conference On Emerging Applications Of Information Technology, 2011, Hal. 253–256.

- [14] B. Santoso, I. Cholissodin, Dan B. D. Setiawan, "Optimasi *K-Means* Untuk Clustering Kinerja Akademik Dosen Menggunakan Algoritme Genetika," *J. Pengemb. Teknol. Inf. Dan Ilmu Komput.*, Vol. 1, No. 12, Hal. 1652–1659, 2017.
- [15] E. A. Sari, "Penerapan Algoritma *K-Means* Untuk Menentukan Tingkat Kesehatan Bayi Dan Balita Pada Kabupaten Dan Kota Di Jawa Tengah," Universitas Dian Nuswantoro, 2015.

BIODATA PENULIS



Pasek Agus Ariawan
 Mahasiswa asal kota Denpasar yang sedang menempuh studi S-2 Konsentrasi Manajemen Sistem Informasi dan Komputer Pada Program Magister Teknik Elektro Universitas Udayana.

LAMPIRAN

Tabel 8. Hasil Pengujian Iterasi Pada Percobaan 1

K	Iterasi	
	<i>K-Means</i> +LOF	<i>K-Means</i>
2	6	7
3	12	7
4	8	5
5	11	24
6	9	16
7	9	16
8	12	20
9	14	21
10	12	11

Tabel 9. Hasil Pengujian Iterasi Pada Percobaan 2

K	Iterasi	
	<i>K-Means</i> +LOF	<i>K-Means</i>
2	8	8
3	9	12
4	6	23
5	16	22
6	10	17
7	8	13
8	14	17
9	12	22
10	20	24

Tabel 10. Hasil Pengujian Iterasi Pada Percobaan 3

K	Iterasi	
	<i>K-Means</i> +LOF	<i>K-Means</i>
2	6	6
3	6	10
4	13	26
5	8	5
6	12	11
7	14	21
8	12	15
9	12	11
10	12	19

Tabel 11. Hasil Pengujian Waktu Pada Percobaan 1

K	Waktu	
	<i>K-Means</i> +LOF	<i>K-Means</i>
2	0.522175	0.538862
3	0.312443	1.12754
4	0.552798	1.313103
5	0.497257	1.630778
6	0.79325	1.091953
7	0.648713	1.861853
8	0.981664	0.898478
9	1.217023	1.575833
10	0.878183	1.079117

Tabel 12. Hasil Pengujian Waktu Pada Percobaan 2

K	Waktu	
	<i>K-Means</i> +LOF	<i>K-Means</i>
2	0.463846	0.615985
3	0.381861	1.180875
4	0.497465	1.415759
5	0.508844	1.817307
6	0.892226	0.747121
7	0.954732	0.979184
8	0.824129	2.019399
9	0.73207	1.806
10	1.017441	0.963188

Tabel 13. Hasil Pengujian Waktu Pada Percobaan 3

K	Waktu	
	<i>K-Means</i> +LOF	<i>K-Means</i>
2	0.307778	0.465295
3	0.454125	0.876056
4	0.390491	0.817293
5	0.46079	0.707313
6	0.393597	0.962908
7	0.799148	2.298508
8	0.589726	1.397508
9	0.645296	1.524811
10	1.424406	1.596745

Tabel 14. Hasil Pengujian SSE Pada Percobaan 1

K	<i>Sum of Square error</i>	
	<i>K-Means+LOF</i>	<i>K-Means</i>
2	2.4999	21.5294
3	1.8247	19.2127
4	1.3301	14.9065
5	1.1455	11.8157
6	1.0433	10.9729
7	0.9054	9.3037
8	0.8732	7.4657
9	0.7659	6.884
10	0.6644	6.0888

Tabel 15. Hasil Pengujian SSE Pada Percobaan 2

K	<i>Sum of Square error</i>	
	<i>K-Means+LOF</i>	<i>K-Means</i>
2	2.501	21.5294
3	1.8244	17.7029
4	1.3289	14.3408
5	1.202	11.9292
6	1.0912	10.3026
7	0.9157	9.1187
8	0.8374	8.6696
9	0.7714	7.0202
10	0.6805	6.617

Tabel 16. Hasil Pengujian SSE Pada Percobaan 3

K	<i>Sum of Square error</i>	
	<i>K-Means+LOF</i>	<i>K-Means</i>
2	2.4999	21.5294
3	1.8242	16.835
4	1.3296	14.6947
5	1.1457	11.9141
6	1.0371	11.547
7	0.9265	9.2075
8	0.859	7.4515
9	0.7608	7.0118
10	0.6842	6.6035