

KOMPARASI ALGORITMA DATA MINING UNTUK CLUSTERING DATA PENDERITA PENYAKIT MENULAR PADA MANUSIA

TUGAS AKHIR

Diajukan Sebagai Salah Satu Syarat
untuk Memperoleh Gelar Sarjana Komputer pada
Program Studi Sistem Informasi

Oleh:

NADIA NURFADILLA

11950320767



FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU
PEKANBARU

2023

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERSETUJUAN

**KOMPARASI ALGORITMA DATA MINING UNTUK
CLUSTERING DATA PENDERITA PENYAKIT MENULAR
PADA MANUSIA**

TUGAS AKHIR

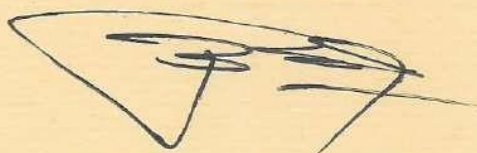
Oleh:

NADIA NURFADILLA

11950320767

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir
di Pekanbaru, pada tanggal 14 Juni 2023

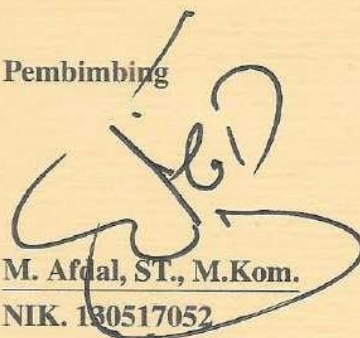
Ketua Program Studi



Eki Saputra, S.Kom., M.Kom.

NIP. 198307162011011008

Pembimbing



M. Afdal, ST., M.Kom.

NIK. 130517052

LEMBAR PENGESAHAN

KOMPARASI ALGORITMA DATA MINING UNTUK CLUSTERING DATA PENDERITA PENYAKIT MENULAR PADA MANUSIA

TUGAS AKHIR

Oleh:

NADIA NURFADILLA


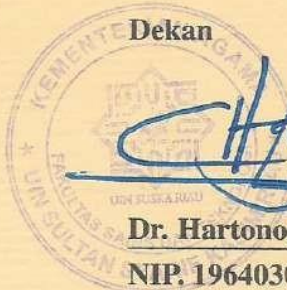
11950320767


Telah dipertahankan di depan sidang dewan penguji
sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer
Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau
di Pekanbaru, pada tanggal 08 Juni 2023

Pekanbaru, 08 Juni 2023

Mengesahkan,

Ketua Program Studi


Dekan

Dr. Hartono, M.Pd.
NIP. 196403011992031003


Eki Saputra, S.Kom., M.Kom.
NIP. 198307162011011008

DEWAN PENGUJI:

Ketua : Siti Monalisa, ST., M.Kom.

Sekretaris : M. Afdal, ST., M.Kom.

Anggota 1 : Inggih Permana, ST., M.Kom.

Anggota 2 : Zarnelly, S.Kom., M.Sc.


Zarnelly

LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum, dengan ketentuan bahwa hak cipta ada pada peneliti. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan atas izin peneliti dan harus dilakukan mengikuti kaedah dan kebiasaan ilmiah serta menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin tertulis dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan dapat meminjamkan Tugas Akhir ini untuk anggotanya dengan mengisi nama, tanda peminjaman dan tanggal pinjam pada *form* peminjaman.

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Dengan menyebut nama Allah yang maha pengasih lagi maha penyayang

Assalamu 'alaikum Warahmatullahi Wabarakaatuh

Alhamdulillah Rabbil 'Alamin, segala puji bagi Allah *Subhanahu Wa Ta'ala*

sebagai bentuk rasa syukur atas segala nikmat yang telah diberikan tanpa ada kekurangan sedikitpun. Shalawat beserta salam tak lupa pula kita ucapkan kepada Nabi Muhammad *Shallallahu 'Alaihi Wa Sallam* dengan mengucapkan *Allahumma Sholli 'ala Sayyidina Muhammad Wa 'ala Ali Sayyidina Muhammad*. Semoga kita semua selalu senantiasa mendapat syafa'at-Nya di dunia maupun di akhirat, *aamiin ya rab-bal 'alaamiin*. Kupersembahkan karya kecil ini sebagai salah satu hadiah istimewa bentuk bakti, rasa terima kasih, dan hormatku kepada orang tuaku tercinta, Ayah dan Mamak.

Ayah dan Mamakku tersayang, terima kasih atas setiap perjuangan, doa, bimbingan, serta dukungan yang kalian berikan kepada saya. Terima kasih atas segala kebaikan dan selalu ada saat keadaan tersulit sekalipun. Terima kasih untuk segala pengorbanan yang kalian lakukan. Sampai kapanpun tiada rasa dan cara yang dapat membalas semuanya. Saya akan selalu mendoakan yang terbaik untuk Ayah dan Mamak agar bahagia dunia dan akhirat, serta diberikan tempat istimewa di sisi-Nya kelak sehingga kita bisa berkumpul kembali bersama-sama di Jannah-Nya.

Terima kasih juga saya ucapkan kepada adik yang sangat saya cintai. Terima kasih untuk segala waktu berharga yang telah dilalui bersama, doa, dan dukungan yang tiada hentinya. Kemudian saya ucapkan terima kasih kepada Bapak dan Ibu Dosen Program Studi Sistem Informasi yang telah mewariskan ilmu yang bermanfaat dan arahan kepada saya untuk menyelesaikan studi di Program Studi Sistem In-formasi ini. Semoga kita semua selalu diberikan kemudahan, rahmat, serta karunia-Nya. *Aamiin*.

Wassalamu 'alaikum Warahmatullahi Wabarakaatuh

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

KATA PENGANTAR

Alhamdulillah Rabbil 'Alamin, bersyukur kehadiran Allah *Subhanahu Wa Ta'ala* atas segala rahmat dan karunia-Nya sehingga peneliti dapat menyelesaikan Tugas Akhir ini dengan baik dan tepat waktu. Shalawat serta salam tidak lupa pula kita ucapkan kepada Nabi Muhammad *Shallallahu 'Alaihi Wa Sallam* dengan mengucapkan *Allahumma Sholli'Ala Sayyidina Muhammad Wa'Ala Ali Sayyidina Muhammad*. Tugas Akhir ini dibuat sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer di Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.

Pada penulisan Tugas Akhir ini, terdapat beberapa pihak yang sudah berkontribusi dan mendukung peneliti baik berupa materi, moril, dan motivasi. Oleh karena itu, peneliti ingin mengucapkan banyak terimakasih kepada:

1. Bapak Prof. Dr. Hairunas, M.Ag sebagai Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Hartono, M.Pd sebagai Dekan Fakultas Sains dan Teknologi.
3. Bapak Eki Saputra, S.Kom., M.Kom sebagai Ketua Program Studi Sistem Informasi sekaligus Dosen Pembimbing sejak Kerja Praktek hingga Tugas Akhir ini.
4. Ibu Siti Monalisa, ST., M.Kom sebagai Sekretaris Program Studi Sistem Informasi sekaligus Ketua Sidang Tugas Akhir peneliti yang telah memberi arahan, saran, serta nasihatnya yang bermanfaat.
5. Bapak Tengku Khairil Ahsyar, S.Kom., M.Kom sebagai Kepala Laboratorium Program Studi Sistem Informasi.
6. Bapak M. Afdal, ST., M.Kom sebagai Dosen Pembimbing Tugas Akhir yang telah banyak meluangkan waktu, memberikan masukan, saran, motivasi, semangat, arahan dan bimbingan yang sangat membantu peneliti dalam pengembangan diri sehingga Tugas Akhir ini dapat diselesaikan dengan baik.
7. Bapak Inggih Permana, ST., M.Kom sebagai Penguji I peneliti yang telah banyak memberikan arahan, masukan, nasihat, serta motivasinya baik dalam penyelesaian Tugas Akhir, maupun juga dalam perkuliahan dan kehidupan sehari-hari. Setiap motivasi yang diberikan akan selalu peneliti ingat dan dijadikan sebagai pelajaran hidup.
8. Ibu Zarnelly, S.Kom., M.Sc sebagai Penguji II peneliti yang telah banyak memberikan arahan, nasihat, masukan, serta motivasinya baik dalam penyelesaian Tugas Akhir, maupun juga dalam perkuliahan dan kehidupan sehari-

hari. Setiap motivasi yang diberikan akan selalu peneliti ingat dan dijadikan sebagai pelajaran hidup.

9. Ibu Fitriani Muttakin, S.Kom., M.Cs sebagai Dosen Pembimbing Akademik peneliti yang telah banyak memberikan arahan, masukan, dan motivasi selama perkuliahan mulai dari Semester 1 hingga Semester 8 ini.
10. Seluruh Bapak dan Ibu Dosen Program Studi Sistem Informasi yang telah banyak memberikan ilmunya kepada peneliti. Semoga ilmu yang diberikan dapat peneliti amalkan dan menjadi amal jariyah.
11. Seluruh Pegawai dan Staff Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau yang telah membantu dan mempermudah proses administrasi selama perkuliahan ini.
12. Kedua orang tua, Ayah dan Mamak tercinta yang selalu mendoakan, memberikan dukungan, memberikan perhatian, kasih sayang dan juga semangat. Terima Kasih atas semua pengorbanan dan kerja keras yang telah kalian lakukan, dengan penuh keikhlasan demi menuju kesuksesan anaknya. Semoga Allah SWT selalu menjaga dan melindungi Ayah dan Mamak dimanapun kalian berada.
13. Adik tunggal M. Ridho Fauzan.T dan seluruh keluarga besar yang telah memberikan semangat, perhatian, motivasi, doa, dan dukungan sehingga peneliti dapat menyelesaikan Tugas Akhir ini.
14. Seluruh Pegawai dan Staff Dinas Kesehatan Kabupaten Pelalawan yang telah membantu dan memberikan dukungan serta arahan kepada peneliti dalam pembuatan Tugas Akhir ini.
15. Seluruh teman-teman Sistem Informasi Angkatan 2019, khususnya Kelas D, terima kasih telah memberikan bantuan dan inspirasi kepada peneliti dalam pembuatan Tugas Akhir ini.
16. Sahabat peneliti, Fitratul Aini, Rahweni Ocviani, Ridho Prasetyo Hanavi, Riskina Saputri, Resi Oktawinda, Suri Yusufi Pratiwi, Windy Amelia Putri dan masih banyak lagi sahabat-sahabat yang telah memberikan dukungan, bantuan, doa, dan inspirasi kepada peneliti dalam pembuatan Tugas Akhir ini.
17. Semua pihak yang tidak dapat peneliti sebutkan satu-persatu, yang terlibat dalam perjuangan penyelesaian pendidikan Strata 1 (S1), yang telah memberikan bantuan, dukungan, dan semangat kepada peneliti baik dalam pengumpulan data maupun dalam penyusunan Tugas Akhir ini.

Semoga segala doa dan dorongan yang telah diberikan selama ini menjadi amal kebajikan dan mendapat balasan setimpal dari Allah *Subhanahu Wa Ta'ala*.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



© Hak cipta milik UIN Suska Riau

Peneliti menyadari bahwa penulisan Tugas Akhir ini masih terdapat banyak kekurangan dan jauh dari kata sempurna. Untuk itu kritik dan saran yang membangun sangat diharapkan demi kesempurnaan Tugas Akhir ini dan semoga laporan ini bermanfaat bagi kita semua. Akhir kata peneliti ucapkan terimakasih.

Pekanbaru, 14 Juni 2023

Peneliti,

NADIA NURFADILLA

NIM. 11950320767



UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

16 Mei 2023

No : 983/LoA/JUTIF/IV/2023
Lampiran : -
Hal : Surat pemberitahuan Bukti Penerimaan Paper

Kepada
Yth

1. Nadia Nurfadilla, UIN Sultan Syarif Kasim Riau
2. M. Afdal, UIN Sultan Syarif Kasim Riau
3. Inggih Permana, UIN Sultan Syarif Kasim Riau
4. Zarnelly, UIN Sultan Syarif Kasim Riau

Jurnal Teknik Informatika (JUTIF) adalah jurnal nasional terdaftar di LIPI dengan P-ISSN : 2723-3863 dan e-ISSN : 2723-3871 serta terakreditasi SINTA 3 berdasarkan Surat Keputusan Direktur Jenderal Pendidikan Tinggi, Riset, Dan Teknologi Kementerian Pendidikan, Kebudayaan, Riset, Dan Teknologi Republik Indonesia Nomor 164/E/KPT/2021.

Kami mengucapkan terima kasih atas pengiriman artikel ilmiah ke **Jurnal Teknik Informatika (JUTIF)** dengan data artikel sebagai berikut:

Judul	<i>COMPARISON OF DATA MINING ALGORITHM FOR CLUSTERING PATIENT DATA HUMAN INFECTIOUS DISEASES</i>
Penulis	<ol style="list-style-type: none">1. Nadia Nurfadilla2. M. Afdal3. Inggih Permana4. Zarnelly

Berdasarkan hasil review, artikel tersebut dinyatakan **DITERIMA** untuk dipublikasikan pada **Jurnal Teknik Informatika (JUTIF) UNIVERSITAS JENDERAL SOEDIRMAN (UNSOED)** Volume 4 Nomor 5 Oktober 2023.

Demikian surat pemberitahuan ini, kami mengucapkan terima kasih atas partisipasi dan kerja sama yang baik.

Chief Editor,



JUTIF
Jurnal Teknik Informatika

Lasmedi Afuan
Dr. Lasmedi Afuan, S.T.,M.Cs
NIP. 19850510 200812 1 002



VOLUME 4 | NOMOR 2 | APRIL 2023



- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

SURAT PERNYATAAN

Saya yang bertandatangan di bawah ini:

Nama : NADIA NURFADILLA
NIM : 11950320767
Tempat/Tgl. Lahir : PELALAWAN / 19 FEBRUARI 2001
Fakultas/Pascasarjana : SAINS DAN TEKNOLOGI
Prodi : SISTEM INFORMASI
Judul Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya*:
KOMPARASI ALGORITMA DATA MINING UNTUK CLUSTERING DATA PENDEKITA
PENYAKIT MENULAK RAPA MANUSIA

Menyatakan dengan sebenar-benarnya bahwa :

1. Penulisan Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya* dengan judul sebagaimana tersebut di atas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya.
3. Oleh karena itu Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya* saya ini, saya nyatakan bebas dari plagiat.
4. Apa bila dikemudian hari terbukti terdapat plagiat dalam penulisan Disertasi/Thesis/Skripsi/(Karya Ilmiah lainnya)* saya tersebut, maka saya bersedia menerima sanksi sesuai peraturan perundang-undangan.

Demikianlah Surat Pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru, 26 Juni 2023
Yang membuat pernyataan



NADIA NURFADILLA
NIM : 11950320767

*pilih salah satu sesuai jenis karya tulis

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

COMPARISON OF DATA MINING ALGORITHM FOR CLUSTERING PATIENT DATA HUMAN INFECTIOUS DISEASES

Nadia Nurfadilla*¹, M. Afdal², Inggih Permana*³, Zarnelly⁴

^{1,2,3,4} Sistem Informasi, Fakultas Sains dan Teknologi, UIN Sultan Syarif Kasim Riau, Indonesia

Email: ¹11950320767@students.uin-suska.ac.id, ²m.afdal@uin-suska.ac.id, ³inggihermana@uin-suska.ac.id, ⁴zarnelly@uin-suska.ac.id

(Naskah masuk: dd mmm yyyy, diterima untuk diterbitkan: dd mmm yyyy)

Abstract

Tuberculosis is known as an infectious disease whose transmission through air intermediaries is caused by the germ *Mycobacterium Tuberculosis*. This disease has become a case that has almost spread throughout the Pelalawan Regency with the number continuing to increase every year so that it is possible to be able to group the areas where this disease spreads. Grouping of tuberculosis data distribution areas using data mining methods in the form of clustering with the data used coming from the Pelalawan Regency Health Office from 2020 to 2022. The data obtained earlier will then be processed using *k-medoids*, *k-means*, and *x-means* algorithms. The beginning of this research was by processing data from each year using these three algorithms. Determination of the most optimal algorithm using DBI or known as the Davies Bouldin Index. The results of the processing of existing indicators are grouped into three sections, namely areas with a high, medium, and low number of cases. From the results of the study, the optimal algorithm in 2020 data is the *k-medoids* algorithms with a DBI value of 0,553 and in 2021 data, the most optimal algorithm is the *k-means* and *x-means* algorithm with similar DBI values of 0,582. Furthermore, the data in 2022 the most optimal algorithms are the *k-means* and *x-means* algorithms because they have the same DBI value, which is 0,510.

Keywords: Clustering, Davies Bouldin Index, Tuberculosis.

KOMPARASI ALGORITMA DATA MINING UNTUK CLUSTERING DATA PENDERITA PENYAKIT MENULAR PADA MANUSIA

Abstrak

Tuberculosis dikenal sebagai penyakit menular yang penularannya melalui perantara udara yang disebabkan oleh bakteri *Mycobacterium Tuberculosis*. Penyakit ini menjadi kasus yang hampir menyebar ke seluruh daerah Kabupaten Pelalawan dengan jumlah yang terus menerus meningkat setiap tahunnya sehingga memungkinkan untuk dapat mengelompokkan daerah yang menjadi tempat penyebaran penyakit ini. Pengelompokan daerah penyebaran *tuberculosis* menggunakan metode *data mining* berupa *clustering* dengan data-data yang digunakan berasal dari Dinas Kesehatan Kabupaten Pelalawan tahun 2020 sampai dengan tahun 2022. Data yang didapatkan tadi kemudian akan diproses atau diolah dengan menggunakan algoritma *k-medoids*, *k-means*, serta *x-means*. Diawalinya penelitian ini dengan memproses data dari tiap tahun dengan menggunakan ketiga algoritma tersebut. Penentuan algoritma paling optimal dengan menggunakan DBI atau yang dikenal dengan *Davies Bouldin Index*. Hasil dari proses pengolahan indikator yang ada dikelompokkan menjadi tiga bagian yaitu daerah yang jumlah kasusnya tinggi, sedang, dan juga rendah. Dari hasil penelitian menghasilkan algoritma yang optimal pada data tahun 2020 yaitu algoritma *k-medoids* dengan nilai DBI yang serupa yaitu 0,553 dan pada data tahun 2021 algoritma yang paling optimal yakni algoritma *k-means* dan *x-means* dengan nilai DBI 0,582. Selanjutnya data pada tahun 2022 algoritma yang paling optimal yakni algoritma *k-means* dan *x-means* dikarenakan mempunyai nilai DBI yang sama yaitu 0,510.

Kata kunci: Clustering, Davies Bouldin Index, Tuberculosis.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. PENDAHULUAN

Angka penderita Penyakit Menular pada Manusia di Indonesia mulai meningkat. Defenisi penyakit menular adalah penyakit yang dapat menjangkiti orang lain dengan ataupun tanpa perantara [1]. *Tuberculosis* atau yang dikenal dengan singkatan TBC merupakan salah satu dari jenis penyakit menular yang mana disebabkan karena adanya bakteri jenis pantogen yang tidak teratasi dengan baik. Selama 77 tahun sejak Negara Indonesia merdeka, sudah banyak dilakukannya upaya untuk mengatasinya, obat serta vaksin dari penyakit ini yang sudah sejak puluhan tahun lalu ditemukan akan tetapi tetap tidak dapat teratasi dengan sempurna [2].

Menurut laporan *World Health Organization* atau disingkat dengan WHO, Negara Indonesia telah terdaftar sebagai salah satu dari 30 negara yang terjangkit TBC tertinggi di dunia dengan menempati urutan peringkat ketiga dari daftar tersebut. Penyakit TBC ini menjadi salah satu penyakit dengan masalah yang ditetapkan oleh WHO menjadi penyakit kedaruratan dunia atau sebutan lainnya *global emergency*. *Tuberculosis* juga merupakan penyakit menular mematikan yang menjadi sebab kematian dengan urutan kedua setelah adanya virus *Covid-19* [3].

Hal yang menjadi latar belakang dilakukannya proses riset ini karena banyak ditemukannya masalah serta kasus terkait *tuberculosis* di provinsi Riau, dimana Riau menjadi salah satu provinsi di urutan kesepuluh yang menduduki peringkat sebagai daerah dengan kasus *tuberculosis* terbanyak. Data yang diambil dalam memproses kasus *tuberculosis* ini merupakan data dari Kabupaten Pelalawan yang terletak di provinsi Riau. Berdasarkan data yang didapat pada Dinas Kesehatan Kabupaten Pelalawan adalah adanya 545 kasus pada tahun 2020, kemudian berkurangnya kasus menjadi 472 kasus di tahun 2021, serta meningkatnya kembali di tahun 2022 sebanyak 626 kasus.

Dari banyaknya kasus permasalahan penyakit *tuberculosis* yang berbahaya untuk masyarakat Indonesia terkhusus masyarakat yang berada di Kabupaten Pelalawan maka perlu adanya upaya pencegahan terhadap penyakit ini dengan dilakukannya analisis *cluster* menggunakan teknik data mining dengan pengelompokan *cluster* tinggi, sedang, juga rendah [4].

Teknik *clustering* memiliki cara kerja dimana ketika didapatkannya populasi atau sekelompok data yang nantinya akan dikelompokkan ke dalam beberapa *cluster* hingga titik data dalam kelompok yang serupa akan lebih mirip dengan titik data lain di kelompok yang sama dibanding yang terdapat pada kelompok lain [5]. Penelitian ini memiliki tujuan berupa untuk diketahuinya pengelompokan *cluster* tinggi, sedang sampai dengan rendah yang kemudian dapat dijadikan bahan evaluasi bagi Dinas

Kesehatan Kabupaten Pelalawan sebagai bentuk upaya yang dapat mencegah penyakit ini berkembang lebih banyak.

Berdasarkan penelitian [6] dikatakan bahwa *k-means* lebih efisien dan efektif dalam mengatasi hasil data dalam ukuran kecil. Dari dataset milik Tunagrahita yang memakai *k-means* memberikan hasil DBI sebanyak 0,161, sedangkan data evaluasi yang menggunakan metode *k-medoids* menghasilkan DBI sebanyak 0,281. Ini dikarenakan *k-medoids* mempunyai performasi komputasi yang lebih tinggi dibandingkan dengan performa komputasi *k-means*. Penelitian [7] ini memiliki tujuan bagi kasus pengelompokan skripsi nantinya yang dapat menjadi tema relevan bagi kelompok yang dibentuk agar dapat mempunyai keakuratan yang tepat. Pengujian dari DBI menggunakan dua algoritma berupa *k-medoids* serta *k-means*. Hasil DBI dari *k-means* yaitu 2,79, sedangkan untuk *k-medoids* yang lebih unggul diperoleh angka 1,56.

Sedangkan penelitian [8] yang memiliki tujuan sebagai pembanding dari analisis hasil dari *x-means* serta *k-means* yang diperoleh dari data masyarakat dan data iris masyarakat miskin di provinsi Aceh tahun 2018. Didapatnya nilai purity dari *x-means* sebanyak 0,88 serta *k-means* sebanyak 0,89 dengan DBI masing-masingnya senilai 0,16. Selama berlangsungnya proses *clustering* dari algoritma *k-means* sendiri membutuhkan waktu selama 0,22 detik dan untuk *x-means* sendiri membutuhkan waktu sebanyak 0,06 detik, sehingga dapat ditarik kesimpulan dari data tersebut bahwa algoritma *k-means* lebih lambat dibandingkan algoritma *x-means*.

Berdasarkan uraian di atas dan di dukung beberapa penelitian sebelumnya, maka pada penelitian ini akan dilakukan perbandingan algoritma *k-means*, *x-means* dan *k-medoids* untuk *clustering* kasus *tuberculosis* dengan menggunakan nilai DBI cluster terbaik sebagai penentu keakuratan algoritma yang digunakan.

2. METODE PENELITIAN

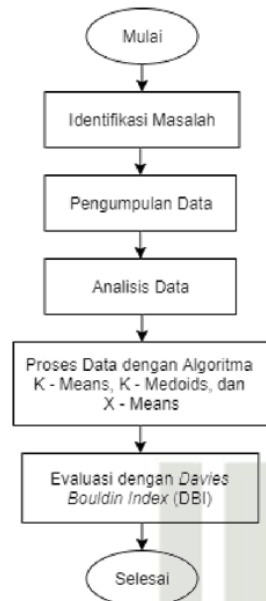
Metode penelitian yang digunakan adalah pendekatan kuantitatif dimana metode penelitian ini dikenal sebagai metode penelitian yang merupakan metode yang terarah secara sistematis dan terstruktur sehingga akan tersusun secara runtut mulai dari awal hingga akhir dilakukannya penelitian ini karena penelitian condong ke arah teknis sistematis berupa analisis angka [9]. Tahap penelitian terlihat pada Gambar 1.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Gambar 1. Metodologi Penelitian

2.1 Identifikasi Masalah

Tahapan pertama dari penelitian ini adalah dilakukannya identifikasi masalah demi mendapatkan informasi, tujuan serta manfaat yang akurat berdasarkan proses penelitian ini [10]. Mengetahui pengelompokan penyebaran kasus *tuberculosis* pada Kabupaten Pelalawan pada tahun 2020 hingga 2022 merupakan salah satu tujuan dari dilaksanakannya penelitian ini. Sedangkan manfaat yang nantinya diperoleh yaitu dapat memberi informasi dan membantu mempermudah perhitungan penyebaran kasus *tuberculosis* kepada masyarakat dan pihak terkait.

2.2 Pengumpulan Data

Dilakukannya wawancara dan observasi sebagai teknik pengumpulan terhadap data kepada pihak instansi bagian Pengendalian dan Pencegahan Penyakit (P2P) yang mana berfokus kepada pengelolaan data terkait *tuberculosis* yang terdapat di Kantor Dinas Kesehatan Kabupaten Pelalawan. Data yang didapatkan berupa data kasus *tuberculosis* tahun 2020 sampai tahun 2022.

2.3 Analisis Data

Proses pengolahan data yang kemudian diubah menjadi sebuah informasi yang diperoleh dari hasil mengidentifikasi ataupun mempelajari data demi kepentingan penelitian disebut dengan analisis data [11]. Pada tahapan ini digunakannya penerapan *data mining* yaitu metode algoritma *clustering*.

2.4 Proses Data menggunakan Algoritma K - Means, K - Medoids, dan X - Means.

Tahap Keempat yaitu memproses data melalui penggunaan algoritma *k-means k-medoids*, dan *x - means*. Dihasilkannya nilai dari pengelompokan

menggunakan tools *RapidMiner Studio* merupakan tujuan dari pengolahan di tahap penelitian ini.

a. Algoritma K-Means

K-means Clustering dikenal sebagai penggunaan analisis metode *data mining* dengan cara memproses pemodelan *unsupervised learning* serta juga menggunakan penggunaan dari cara pengelompokan beberapa partisi atau bagian. Langkah algoritma metode *k-means* [12]:

1. Menentukan jumlah pada kelompok (*cluster*)
2. Pilih pusat *cluster*
3. Hitung pusat kelompok (*centroid*) / rata - rata dari data yang ada masing masing kelompok. Lokasi *centroid* menggunakan rata - rata (*mean*) dari semua nilai data di setiap fiturnya.
4. Alokasikan kembali setiap data dari *centroid* yang terdekat
5. Kembali ke langkah 3, bila masih ditemukan adanya transfer data dari satu *cluster* ke *cluster* yang lain, ataupun jika terdapat perubahan yang didapat pada nilai *centroid* di atas nilai *threshold* yang ditentukan.
6. Digunakannya *Euclidean Distance* [13] untuk dapat menghitung jarak semua data kepada setiap titik pusat *cluster* dengan menggunakan persamaan (1).

$$D_{ij} = \sqrt{(X_{1i} - X_{1j})^2 \dots + (X_{ki} - X_{kj})^2} \quad (1)$$

Dimana :

D(ij) = jarak data nilai (i) kepusat *cluster* (j)

X_{ki} = data nilai (i) pada atribut data ke (k)

X_{kj} = titik pusat (j) pada atribut ke (k)

b. Algoritma K-Medoids

Penggunaan objek pada setiap kelompok objek nantinya sebagai perwakilan sebuah *cluster* di algoritma ini. Objek yang kemudian terpilih akan menjadi perwakilan sebuah *cluster* yang kemudian dikenal dengan nama *medoids*. Pembangunan *cluster* dengan memperhitungkan adanya kedekatan antara objek *non medoids* dengan objek *medoids* [14]:

1. Sebanyak jumlah *cluster* (k) akan diinisialisasi pusat *cluster*.
2. Digunakannya persamaan pada ukuran jarak *euclidean distance* yang ada di persamaan (2.1) untuk dialokasikan pada tiap objek ke *cluster* terdekat.
3. Dipilihnya secara acak objek masing-masing *cluster* selayaknya calon *medoids* yang baru.
4. Dihitungnya jarak pada tiap objek yang terdapat di masing-masing *cluster* bersama dengan calon *medoids* yang baru.
5. Dihitungnya total simpangan (S) dengan cara hitung nilai total dari *distance* baru - total *distance* lama. Jika data yang didapat adalah S

6. < 0 , maka tukar objek dengan data *cluster* demi menghasilkan sekumpulan k objek baru sebagai *medoids*.
7. Kemudian ulang kembali langkah 3 sampai 5 sampai tidak adanya perubahan *medoids*, yang kemudian akan didapatkan *cluster* beserta anggota *cluster* masing-masing.

c. Algoritma *X-Means*

Kemudian adanya pengembangan dari algoritma *k-means* dikenal sebagai algoritma *x-means* yang mana *x-means* dihasilkan dari kelompok operasi berulang sehingga didapatnya nilai *Bayesian Information Criterion* atau disingkat *BIC* hingga akhir eksekusi. Berikut merupakan langkah dari algoritma *x-means* [15]:

1. Ditentukannya range dari *cluster* $K(Kmin, Kmax)$.
2. Dilakukannya inisiasi nilai $K = Kmin$.
3. Dijalankannya *x-means* sampai dengan konvergen (tidak ada perubahan terhadap anggota *cluster* ataupun letak *cluster*nya).
4. Memperbaiki struktur, langkah ini dimulai dengan memecah setiap centroid hasil langkah 3 menjadi dua children dalam arah yang berlawanan di sepanjang vektor yang dipilih secara acak. Setelah itu menjalankan *x-means* secara lokal di dalam setiap *cluster* untuk dua *cluster*. Keputusan masing-masing pusat *cluster* sendiri dengan membandingkan nilai-nilai *BIC*.
5. Dilakukannya pembaruan dari nilai K , apabila $K > Kmax$ maka proses akan berhenti dan akan melaporkan struktur terbaik yang telah ditemukan maka akan kembali ke langkah 3.

2.5 Evaluasi dengan DBI

Pada tahap ini dilakukan evaluasi hasil tiga algoritma tersebut. Evaluasi ini dilakukan untuk melihat kualitas *cluster* yang ada [16] Pengevaluasian dilaksanakan dengan menggunakan *Davies Bouldin Index* atau yang lebih dikenal dan

disingkat dengan DBI. DBI merupakan salah satu metode analisis untuk mengukur kualitas *cluster* di setiap *clustering* [17]. DBI juga dikenal sebagai fungsi dari rasio sekelompok distribusi *cluster* selayaknya pemisah antara *cluster-cluster* [18]. Nilai kohesi dan separasi dikenal sebagai bagian dari jenis pendekatan dipengujian nilai DBI.

3. HASIL DAN PEMBAHASAN

Terdapat 4 pembagian dari hasil pembahasan penelitian, yaitu adanya Dataset, *Clustering*, Evaluasi dan Hasil Penentuan Algoritma Terbaik.

3.1 Dataset

Dataset Penelitian ini menggunakan data kasus *tuberculosis* di Kabupaten Pelalawan dan disimpan dalam bentuk file *Microsoft Excel*. Data yang digunakan dimulai dari tahun 2020 sampai 2022. Variabel yang digunakan dan penjelasannya terlihat pada Tabel 1.

Tabel 1. Variabel yang digunakan

Variabel	Inisial	Penjelasan
Jumlah Kasus	JK	Total keseluruhan kasus pada suatu kecamatan
Usia Produktif	UP	Usia dari rentang 15 – 64 tahun
Non Produktif	NP	Usia dibawah 15 tahun dan diatas dari 64 tahun
Laki - Laki	LK	Jenis kelamin laki laki
Perempuan	PR	Jenis kelamin perempuan
Pelajar/Mahasiswa	PM	Status pasien yang masih pelajar/mahasiswa
Tidak Bekerja	TK	Pasien yang tidak bekerja
Bekerja	K	Pasien yang bekerja
Tidak Diketahui	TD	Pasien yang tidak diketahui pekerjaannya
TBC Extraparu	TE	Pasien penderita <i>tuberculosis</i> jenis extraparu
TBC Paru	TP	Pasien penderita <i>tuberculosis</i> jenis paru

Pada tahun 2020 kasus *tuberculosis* mencapai 545 kasus di Kabupaten Pelalawan. Tabel 2 merupakan data kasus *tuberculosis* pada tahun 2020.

Tabel 2. Data Tahun 2020

Kecamatan	JK	UP	NP	LK	PR	PM	TK	K	TD	TE	TP
Bandar Petalangan	1	1	0	1	0	0	0	0	0	0	1
Bandar Seikijang	29	29	0	13	16	2	7	9	11	0	29
Bunut	20	20	0	11	9	2	5	9	4	0	20
Kerumutan	21	20	1	15	6	1	1	9	4	1	20
Kuala Kampar	20	20	0	12	8	2	5	8	5	2	18
Langgam	56	52	4	35	21	2	13	16	25	5	51
Pangkalan Kerinci	202	187	15	126	76	10	39	61	92	7	195

Kecamatan	JK	UP	NP	LK	PR	PM	TK	K	TD	TE	TP
Pangkalan Kuras	57	56	1	37	20	2	11	22	22	1	56
Pangkalan Lesung	40	40	0	25	15	2	7	12	19	1	39
Pelalawan	41	40	1	22	19	1	12	12	16	3	38
Teluk Meranti	13	12	1	10	3	1	1	4	7	0	13
Ukui	45	45	0	25	20	3	8	6	28	3	42
Total Kasus	545										

Pada tahun 2021 terjadi penurunan kasus *tuberculosis* di Kabupaten Pelalawan. Tabel 3 merupakan data kasus *tuberculosis* pada tahun 2021.

Tabel 3. Data Tahun 2021

Kecamatan	JK	UP	NP	LK	PR	PM	TK	K	TD	TE	TP
Bandar Petalangan	5	5	0	3	2	1	0	3	1	1	4
Bandar Seikijang	28	28	0	14	14	0	5	12	11	1	27
Bunut	12	12	0	8	4	0	2	5	5	0	12
Kerumutan	18	17	1	8	10	1	4	7	6	0	18
Kuala Kampar	20	19	1	10	10	0	8	9	3	0	20
Langgam	50	48	2	29	21	3	7	24	16	1	49
Pangkalan Kerinci	168	163	5	111	57	5	31	83	49	4	164
Pangkalan Kuras	57	56	1	37	20	2	13	26	16	3	54
Pangkalan Lesung	36	33	3	24	12	0	9	16	11	0	36
Pelalawan	28	26	2	22	6	2	4	14	8	1	27
Teluk Meranti	11	9	2	7	4	1	1	1	8	0	11
Ukui	39	37	2	24	15	3	5	24	7	3	36
Total Kasus	472										

Pada tahun 2022 kasus *tuberculosis* di Kabupaten Pelalawan meningkat drastis hingga mencapai 625 kasus. Tabel 4 merupakan data tahun 2022.

Tabel 4. Data Tahun 2022

Kecamatan	JK	UP	NP	LK	PR	PM	TK	K	TD	TE	TP
Bandar Petalangan	27	27	0	18	9	11	5	16	8	2	25
Bandar Seikijang	24	24	0	17	7	15	7	16	4	3	21
Bunut	14	13	1	10	4	12	0	6	8	0	14
Kerumutan	31	29	2	20	11	22	6	24	6	0	31
Kuala Kampar	29	27	2	19	10	2	6	18	5	1	6
Langgam	39	35	4	24	15	36	2	26	11	0	39
Pangkalan Kerinci	183	177	6	116	67	150	37	83	70	9	174
Pangkalan Kuras	96	92	4	57	39	3	17	57	34	4	92
Pangkalan Lesung	50	48	2	31	18	2	6	31	17	2	47
Pelalawan	38	37	1	24	14	32	0	17	21	0	38
Teluk Meranti	27	24	3	19	7	22	3	18	8	1	26
Ukui	67	60	7	43	24	61	8	41	16	8	59
Total Kasus	625										

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

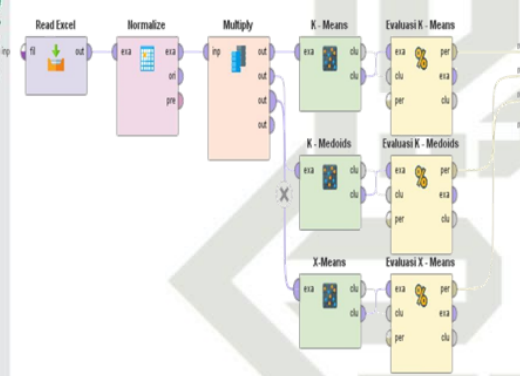


Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3.2 Clustering

Dari data yang telah tersedia kemudian dilanjutkan pada proses pemodelan dimana akan adanya pengolahan data menjadi sebuah *cluster* menggunakan algoritma *k-means*, *k-medoids*, dan *x-means* yang sesuai tahapan dari aturan algoritma itu. Berdasarkan data yang didapat dari tiga tahun sebelumnya akan dilakukan proses perbandingan algoritma. Data kemudian di proses sesuai dengan urutan hingga nantinya akan didapatkan hasil yang diinginkan agar mendapatkan hasil optimal dan terbaik. Penggunaan *tools* dalam mengolah data ini adalah dengan *RapidMiner Studio* yang berfungsi sebagai alat yang dapat menganalisis *data mining*, *text mining* serta analisis prediksi [19]. Gambar 2 merupakan proses *clustering* data pada *RapidMiner Studio*.



Gambar 2. Proses di RapidMiner Studio

Hasil pengolahan *clustering* data dari algoritma *k-means*, *k-medoids*, dan *x-means* dari tahun 2020 sampai 2022 terlihat pada Tabel 5.

Tabel 5. Hasil *clustering* Data di RapidMiner Studio

Tahun	K	K - Means	K - Medoids	X - Means
2020	0	6	6	4
	1	1	1	1
	2	5	5	7
2021	0	1	8	1
	1	8	1	8
	2	3	3	3
2022	0	1	1	1
	1	9	9	9
	2	2	2	2

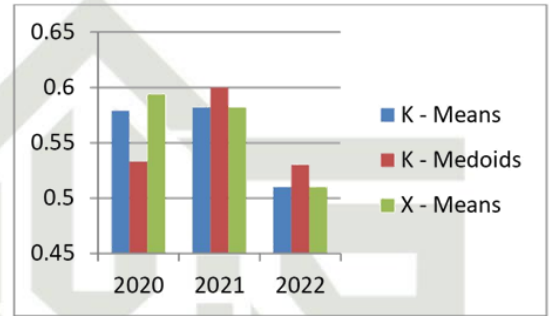
3.3 Evaluasi

Data yang telah diproses pemodelan dengan penggunaan teknik *clustering* lalu dievaluasi [20]. Hasil Evaluasi dari algoritma *k-means*, *k-medoids*, dan *x-means* dari tahun 2020 sampai 2022 terlihat pada Tabel 6.

Tabel 6. Hasil Evaluasi Data di RapidMiner Studio

Tahun	K - Means	K - Medoids	X - Means
2020	0,579	0,553	0,594
2021	0,582	0,600	0,582
2022	0,510	0,530	0,510

Penggunaan tahap ini dikhususkan untuk pengevaluasian kinerja dari algoritma-algoritma yang menggunakan teknik *data mining* untuk mengidentifikasi algoritma yang tepat dalam penerapan pada dataset. Teknik evaluasi yang digunakan yaitu *Davies Bouldin Index* (DBI). Perbandingan nilai DBI pada data tiap tahunnya terlihat pada gambar 3.



Gambar 4. Perbandingan Nilai DBI

Nilai DBI terendah pada data tahun 2020 yaitu algoritma *k-medoids* memiliki nilai DBI yang sama yaitu 0,533, selanjutnya pada data tahun 2021 nilai DBI terendah berada pada algoritma *k-means* dan *x-means* dengan nilai DBI 0,582, dan pada data tahun 2022 nilai DBI terendah berada pada algoritma *k-means* dan *x-means* dengan nilai yang sama 0,510.

3.4 Hasil Penentuan Algoritma Terbaik

Hasil dari perbandingan algoritma dari tahun 2020 sampai 2022 memiliki anggota *cluster* yang berbeda, berikut anggota *cluster* dari algoritma terbaik pada tahunnya.

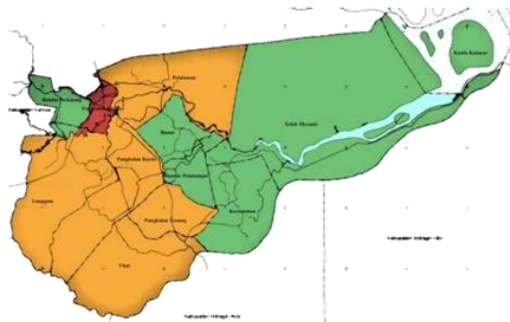
a. Anggota Cluster Tahun 2020

Pada tahun 2020 algoritma dengan nilai DBI terendah yakni algoritma *k-medoids* anggota *cluster* dapat ditemukan di Tabel 7.

Tabel 7. Anggota Cluster Data Tahun 2020

Cluster	Status	Anggota
0	Rendah	Bandar Petalangan, Bandar Seikijang, Bunut, Kerumutan, Kuala Kampar, Teluk Meranti
1	Tinggi	Pangkalan Kerinci
2	Sedang	Langgam, Pangkalan Kuras, Pangkalan Lesung, Pelalawan, Ukui

Tabel 7 menjelaskan terkait kecamatan yang telah menjadi anggota dari bagian *cluster-cluster* dimana terdapat 6 anggota di *cluster* 0, 1 anggota di *cluster* 1, serta 5 anggota di *cluster* 3. Apabila disajikan dalam bentuk peta wilayah maka visualnya terlihat seperti di Gambar 5.



Gambar 5. Peta Wilayah Cluster Tahun 2020

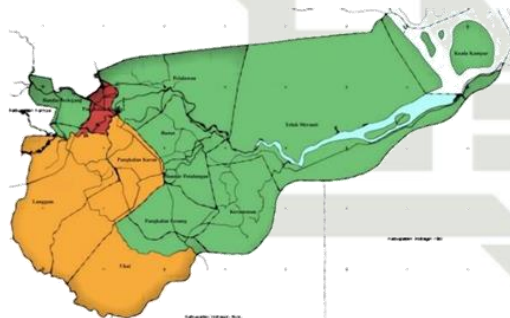
b. Anggota Cluster Tahun 2021

Pada tahun 2021 algoritma dengan nilai DBI terendah yakni algoritma *k-means* dan *x-means* anggota cluster dapat ditemukan di Tabel 8.

Tabel 8. Anggota Cluster Data Tahun 2021

Cluster	Status	Anggota
0	Tinggi	Pangkalan Kerinci
1	Rendah	Bunut, Kuala Kampar, Bandar Seikijang, Teluk Mearnti, Bandar Petalangan, Kerumutan, Pelalawan, dan Pangkalan Lesung
2	Sedang	Langgam, Pangkalan Kuras, Ukui

Tabel 8 menjelaskan terkait kecamatan yang telah menjadi anggota dari *cluster-cluster* yang terbentuk, dimana terdapat 1 anggota di *cluster* 0, 8 anggota di *cluster* 1, serta 3 anggota di *cluster* 3. Apabila disajikan dalam bentuk peta wilayah maka visualnya terlihat seperti di Gambar 6



Gambar 6. Peta Wilayah Cluster Tahun 2021

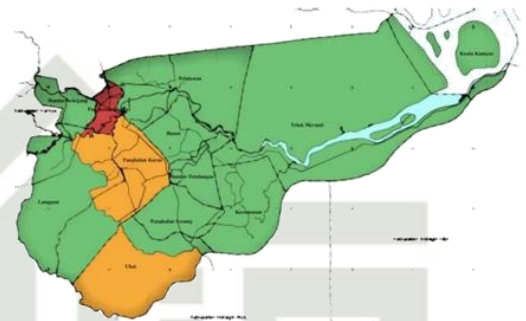
c. Anggota Cluster Tahun 2022

Pada tahun 2022 algoritma dengan nilai DBI terendah yakni algoritma *k-means* dan *x-means* anggota cluster dapat ditemukan di Tabel 9.

Tabel 9. Anggota Cluster Data Tahun 2022

Cluster	Status	Anggota
0	Tinggi	Pangkalan Kerinci
1	Rendah	Bandar Petalangan, Bandar Seikijang, Bunut, Kerumutan, Kuala Kampar, Langgam, Pelalawan, Pangkalan Lesung, Teluk Meranti
2	Sedang	Pangkalan Kuras, Ukui

Tabel 8 menjelaskan terkait kecamatan yang telah menjadi anggota dari *cluster-cluster* yang terbentuk, dimana terdapat 1 anggota di *cluster* 0, 9 anggota di *cluster* 1, serta 2 anggota di *cluster* 2. Apabila disajikan dalam bentuk peta wilayah maka visualnya terlihat seperti di Gambar 7.



Gambar 7. Peta Wilayah Cluster Tahun 2022.

4. DISKUSI

Diterapkannya *data mining* sebagai penentu dalam pola penyebaran penyakit *tuberculosis* yang terdapat di Kabupaten Pelalawan melalui metode *clustering* dan bantuan *tools software Rapidminer Studio* untuk mampu menentukan daerah yang telah terjangkit penyakit *tuberculosis*. Berdasarkan proses-proses yang ada dan telah dilaksanakan maka didapatkan kesimpulan bahwa tingkat kategori penyebaran penyakit *Tuberculosis* ini terdiri atas kategori tinggi, sedang, serta rendah. Nilai DBI terendah terdapat pada algoritma *k-means* dan *x-means* pada data tahun 2022 dengan nilai 0,510. Menurut penelitian ini maka hasil dari *clustering* data setiap tahunnya memiliki tingkat keakuratan yang cukup dan didukung dengan adanya peta wilayah untuk membantu pemerintah dalam penanganan kasus *tuberculosis* di Kabupaten Pelalawan. Hasil dari penelitian ini dan meninjau dari penelitian terdahulu [21] terdapat perbedaan dalam jumlah atribut yang digunakan dan data yang diolah, dimana pada penelitian ini data diolah persetiap tahunnya dan menunjukkan dengan peta wilayah agar terlihat perbedaan hasil dari tiap tahunnya.

5. KESIMPULAN

Didapatnya kesimpulan berdasarkan penelitian yang telah dilakukan dimana diantaranya adalah metode *k-means*, *k-medoids* dan *x-means clustering* telah berhasil diimplementasikan untuk mengelompokkan pola penyebaran kasus *tuberculosis* di Kabupaten Pelalawan. Berdasarkan pelaksanaan *clustering*, tidak ada perbedaan yang jauh antara algoritma *k-means*, *k-medoids*, dan *x-means*. Jumlah *cluster* yang dibentuk pada penelitian ini menggunakan nilai $K=3$.

Data tiap tahun memiliki algoritma optimal yang berbeda. Pada tahun 2020 algoritma yang paling optimal ialah *k-medoids* dengan nilai DBI yang serupa yaitu 0,553. Kecamatan dengan status tinggi yaitu Pangkalan Kerinci, status sedang beranggotakan Langgam, Pangkalan Kuras, Pangkalan Lesung, Pelalawan, Ukui dan status rendah memiliki anggota Bandar Petalangan, Bandar Seikijang, Bunut, Kerumutan, Kuala Kampar, Teluk Meranti.

Data pada tahun 2021 algoritma yang paling optimal yaitu *k-medoids* dengan nilai DBI 0,582 dengan hasil anggota kecamatan dengan status tinggi yaitu Pangkalan Kerinci, status sedang beranggotakan Langgam, Pangkalan Kuras, Ukui sedangkan status rendah beranggotakan Bandar Petalangan, Bandar Seikijang, Bunut, Kerumutan, Kuala Kampar, Pangkalan Lesung, Pelalawan, dan Teluk Meranti.

Sedangkan pada data tahun 2022 algoritma yang paling optimal ialah *k-means* dan *x-means* yang juga memiliki nilai DBI yang sama yaitu 0,510 kecamatan dengan status tinggi yaitu Pangkalan Kerinci, status sedang Pangkalan Kuras dan Ukui sedangkan status rendah yaitu Bunut, Kuala Kampar, Bandar Seikijang, Langgam, Teluk Meranti, Pelalawan, Bandar Petalangan, Kerumutan, Pangkalan Lesung.

DAFTAR PUSTAKA

- [1] V. T. Hulu *et al.*, *Epidemiologi Penyakit Menular: Riwayat, Penularan dan Pencegahan*. Yayasan Kita Menulis, 2020.
- [2] Adt/Ning, "Melalui Kegiatan INA – TIME 2022 Ke-4, Menkes Budi Minta 90% Penderita TBC Dapat Terdeteksi di Tahun 2024," *p2p.kemkes.go.id*, 2022. <http://p2p.kemkes.go.id/melalui-ina-time-2022-ke-4-menkes-budi-minta-90-penderita-tbc-dapat-terdeteksi-di-tahun-2024/>.
- [3] WHO/Indonesia, "Hari Tuberkulosis Sedunia 2022," 2022. <https://www.who.int/indonesia/news/campaign/tb-day-2022>.
- [4] D. N. Batubara, A. P. Windarto, D. Hartama, and H. Satria, "Analisis Metode K-MEANS Pada Pengelompokan Keberadaan Area Resapan Air Menurut Provinsi," *Semin. Nas. Sains Teknol. Inf.*, no. x, pp. 345–349, 2019, [Online]. Available: <https://prosiding.seminar-id.com/index.php/sensasi/article/view/324> (05 Juni 2020).
- [5] W. A. Wahyuni and S. Saepudin, "Penerapan Data Mining Clustering Untuk Mengelompokkan Berbagai Jenis Merk Mesin Cuci," *Semin. Nas. Sist. ...*, pp. 306–

313, 2021, [Online]. Available: <https://sisematik.nusaputra.ac.id/index.php/sisematik/article/view/35%0Ahttps://sisematik.nusaputra.ac.id/index.php/sisematik/article/download/35/31>.

- [6] F. Harahap, "Perbandingan Algoritma K Means dan K Medoids Untuk Clustering Kelas Siswa Tunagrahita," *TIN Terap. Inform. Nusant.*, vol. 2, no. 4, pp. 191–197, 2021.
- [7] S. Ramadhani, D. Azzahra, U. I. Negeri, and S. S. Kasim, "Comparison of K-Means and K-Medoids Algorithms in Text Mining based on Davies Bouldin Index Testing for Classification of Student 's Thesis," vol. x, no. x, pp. 24–33, 2022.
- [8] B. Bustami, R. Mahara, H. Ahmadian, S. Wahyuni, and K. AR, "Analisis Clustering Penduduk Miskin Di Provinsi Aceh Menggunakan Algoritma K-Means Dan X-Means," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 5, no. 1, pp. 26–35, 2022, doi: 10.32672/jnkti.v5i1.3961.
- [9] I. N. Sari *et al.*, *Metode penelitian kualitatif*. UNISMA PRESS, 2022.
- [10] A. Szent-Gyorgyi, "IDENTIFIKASI PENELITIAN DAN PERUMUSAN MASALAH PENELITIAN," *Pengantar Metodol. Penelit. Pendekatan Multidisipliner*, p. 19, 2022.
- [11] F. Hasyim and M. Muafi, "Implementasi Data Mining Dalam Menentukan Strategi Promosi Program KB Menggunakan Algoritma K-Means Clustering," *COREAI J. Kecerdasan Buatan ...*, vol. 3, no. 1, pp. 110–119, 2022, [Online]. Available: <https://ejournal.unuja.ac.id/index.php/coreai/article/view/4292>.
- [12] E. Muningsih, I. Maryani, and V. R. Handayani, "Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa," *J. Sains dan Manaj.*, vol. 9, no. 1, pp. 95–100, 2021, [Online]. Available: www.bps.go.id.
- [13] A. Rohmah, F. Sembiring, and ..., "Implementasi Algoritma K-Means Clustering Analisis Untuk Menentukan Hambatan Pembelajaran Daring (Studi Kasus: Smk Yaspim ...)," ... *Sist. Inf. dan ...*, pp. 290–298, 2021, [Online]. Available: <https://sisematik.nusaputra.ac.id/index.php/sisematik/article/view/32>.
- [14] H. D. Tampubolon, S. Suhada, M. Safii, S. Solikhun, and D. Suhendro, "Penerapan Algoritma K-Means dan K-Medoids Clustering untuk Mengelompokkan Tindak Kriminalitas Berdasarkan Provinsi," *J. Ilmu*

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

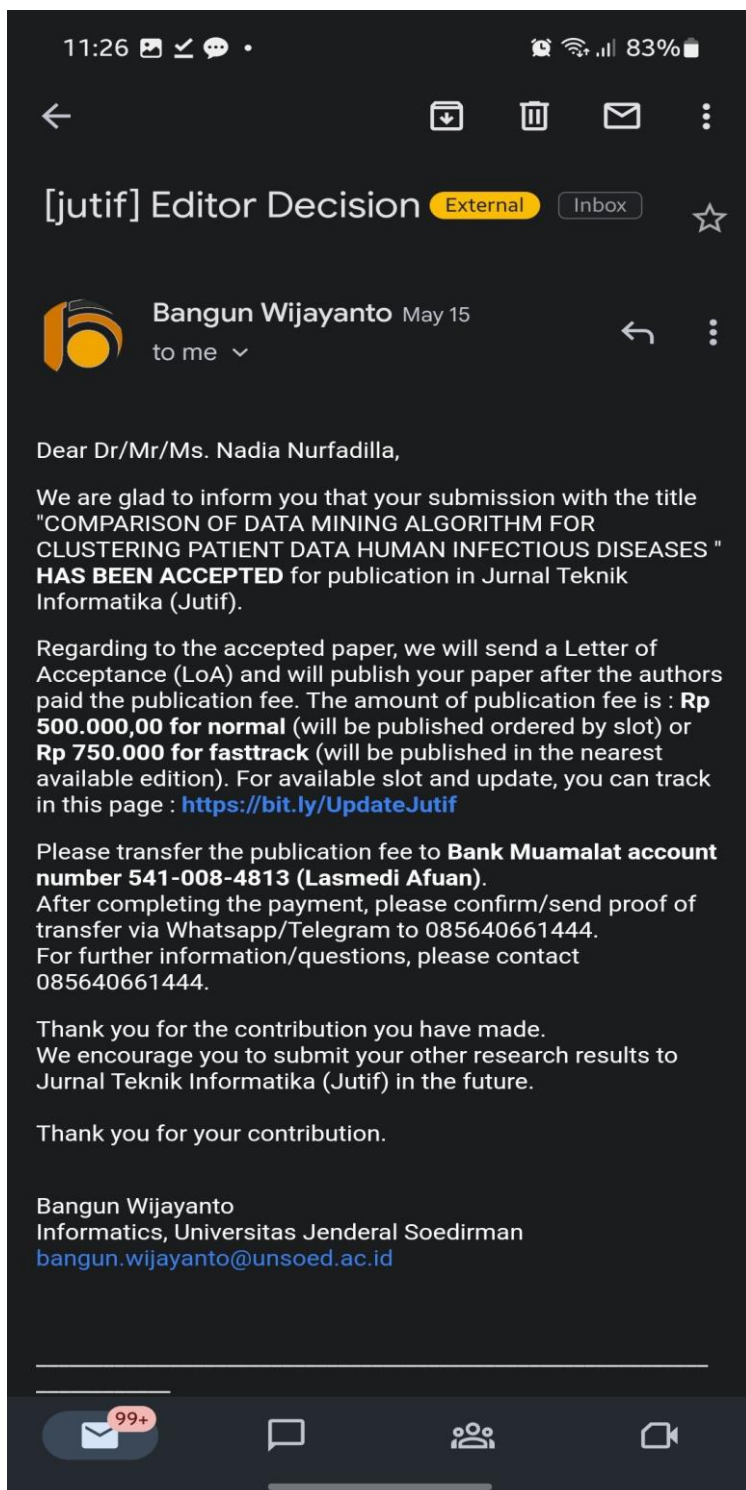
- Komput. dan Teknol.*, vol. 2, no. 2, pp. 6–12, 2021, doi: 10.35960/ikomti.v2i2.703.
- [15] A. A. Riza and D. R. S. Saputro, “Clustering Data Numerik Menggunakan Algoritme X-Means,” *UNEJ e-Proceeding*, 2022, [Online]. Available: <https://jurnal.unej.ac.id/index.php/prosiding/article/view/33491%0Ahttps://jurnal.unej.ac.id/index.php/prosiding/article/download/33491/11655>.
- [16] A. F. Solikin, Kusriani, and F. W. Wibowo, “Evaluasi Cluster Data Interkomparasi Anak Timbangan Dengan Algoritma Self Organizing Maps Cluster Evaluation Weighing Intercomparison Data with Self Organizing Maps Algorithm,” vol. 11, no. 2, pp. 208–219, 2021.
- [17] D. A. Saidah, R. Santoso, and T. Widiharih, “Pengelompokan Provinsi Di Indonesia Berdasarkan Indikator Kesehatan Lingkungan Menggunakan Metode Partitioning Around Medoids Dengan Validasi Indeks Internal,” *J. Gaussian*, vol. 11, no. 2, pp. 302–312, 2022, doi: 10.14710/j.gauss.v11i2.35478.
- [18] T. A. Munandar, “Penerapan Algoritma Clustering Untuk Pengelompokan Tingkat Kemiskinan Provinsi Banten,” *JSiI (Jurnal Sist. Informasi)*, vol. 9, no. 2, pp. 109–114, 2022, doi: 10.30656/jsii.v9i2.5099.
- [19] G. T. Enggiel, H. D. Purnomo, and D. K. Salatiga, “APPLICATION OF K-MEANS METHOD IN THE SPREAD OF POSITIVE CASES OF COVID-19 IN SALATIGA CITY Penerapan Metode K-Means Pada Penyebaran Kasus Positif Covid-19,” vol. 3, no. 5, pp. 1323–1328, 2022.
- [20] J. Ilmiah and W. Pendidikan, “Penerapan K-Means Clustering dalam Pengelompokan Kasus Tuberkulosis di Provinsi Jawa Barat,” *J. Ilm. Wahana Pendidik.*, vol. 8, no. September, pp. 219–227, 2022.
- [21] A. J. Wahidin and D. I. Sensuse, “Perbandingan Algoritma K-Means, X-Means Dan K-Medoids Untuk Klasterisasi Awak Kabin Lion Air,” *J. ICT Inf. Commun. Technol.*, vol. 20, no. 2, pp. 298–302, 2021, doi: 10.36054/jict-ikmi.v20i2.387.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

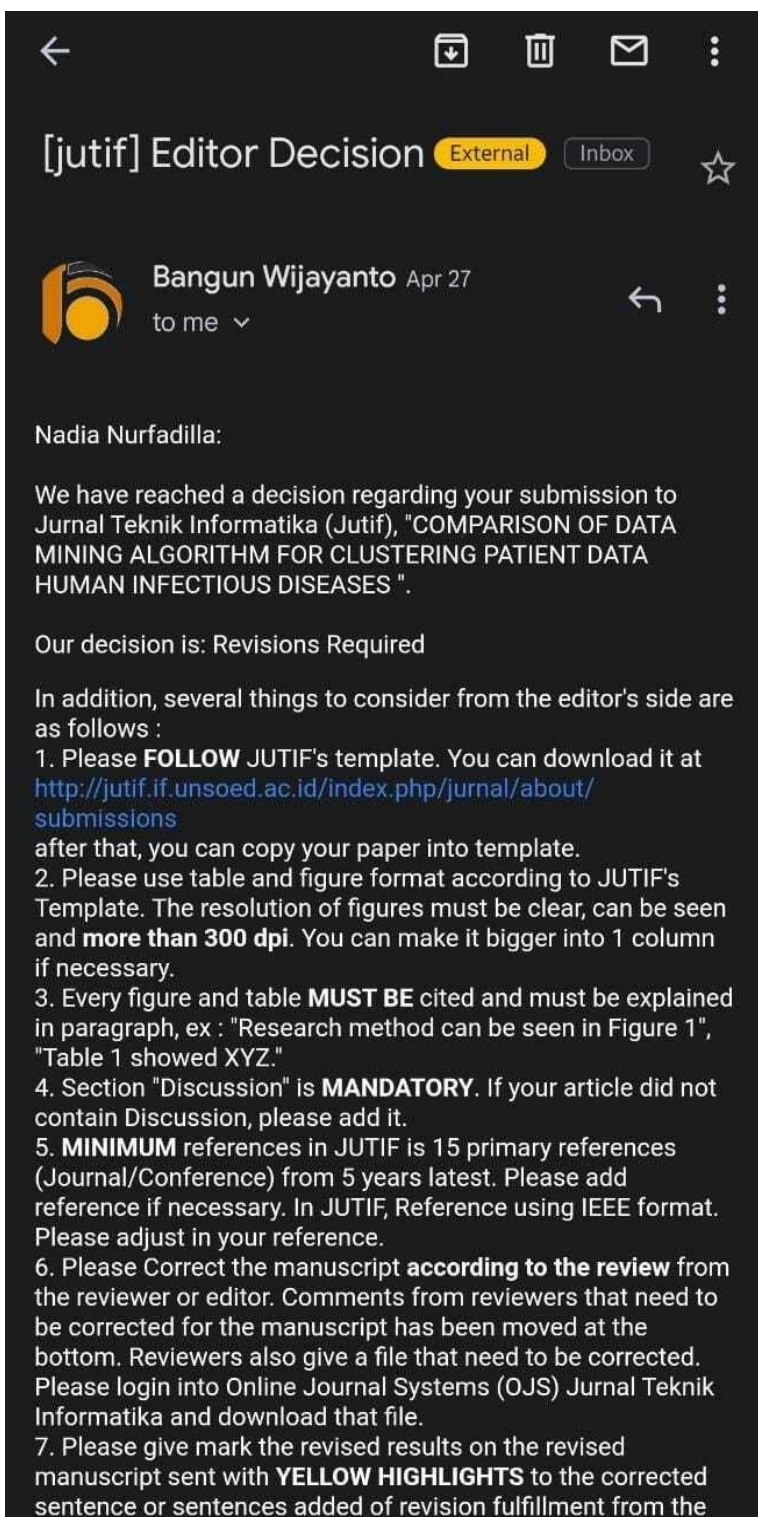
LAMPIRAN A

BUKTI ACCAPTANCE PAPER



LAMPIRAN B

REVISI PAPER



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



LAMPIRAN C DOKUMENTASI



© Hak cipta

Farif Kasim Riau

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LAMPIRAN D

SURAT PENELITIAN



PEMERINTAH KABUPATEN PELALAWAN
DINAS PENANAMAN MODAL
DAN PELAYANAN TERPADU SATU PINTU

Komplek Perkantoran Bhakti Praja Email : dpmptsp.pelalawan@gmail.com
 Telp/Fax : 0761-95992. Telp : 0761-95991
 PANGKALAN KERINCI

REKOMENDASI

Nomor : 504/DPMPTSP/2022/0167

PELAKSANAAN KEGIATAN RISET / PRA RISET DAN PENGUMPULAN DATA UNTUK BAHAN PENELITIAN

Berdasarkan Peraturan Bupati Pelalawan Nomor 16 Tahun 2019 Tentang Pendelegasian Wewenang Menandatangani Perizinan dan Non Perizinan serta Penyelenggaraan Pelayanan kepada Kepala Dinas Penanaman Modal dan Pelayanan Terpadu Satu Pintu Kabupaten Pelalawan dan Surat dari DINAS PENANAMAN MODAL DAN PELAYANAN TERPADU SATU PINTU PROVINSI RIAU Nomor : Nomor : 503/DPMPTSP/NON IZIN-RISET/0 dengan ini memberikan rekomendasi kepada :

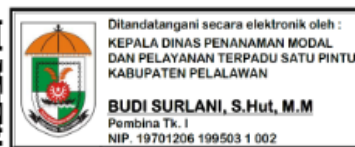
1. Nama : **NADIA NURFADILLA**
2. NIM / KTP : 11950320767
3. Program Studi : SISTEM INFORMASI
4. Jenjang : S1
5. Alamat : JL. JAMBU GG. ASY SUKRON
6. Judul Penelitian : ANALISA PENYEBARAN PENYAKIT MENULAR PADA MANUSIA DENGAN ALGORITMA K-MEANS DAN FP-GROWTH
7. Lokasi Penelitian : DINAS KESEHATAN KABUPATEN PELALAWAN

Dengan ketentuan sebagai berikut :

1. Tidak melakukan kegiatan yang menyimpang dari ketentuan yang telah ditetapkan yang tidak ada hubungan dengan Riset / Pra Riset dan pengumpulan data ini.
2. Pelaksanaan riset ini berlangsung selama 6 (enam) bulan terhitung tanggal rekomendasi ini dibuat.

Demikian rekomendasi ini diberikan agar dapat dipergunakan sebagaimana mestinya dan kepada pihak yang terkait diharapkan untuk dapat memberikan kemudahan dan membantu kelancaran kegiatan penelitian dan pengumpulan data ini dan terima kasih.

Dikeluarkan di Pangkalan Kerinci
 Pada tanggal 24 Oktober 2022



Tembusan :

1. Dinas Kesehatan Kabupaten Pelalawan
2. Kepala Badan Kesatuan Bangsa dan Politik Kabupaten Pelalawan
3. Dekan Fakultas Sains dan Teknologi UIN Suska Riau di Pekanbaru
4. Yang Bersangkutan

DAFTAR RIWAYAT HIDUP

Nadia Nurfadilla adalah nama peneliti Tugas Akhir ini. Peneliti lahir di Pekanbaru pada tanggal 19 Februari 2001. Peneliti adalah anak dari Bapak Abdul Kadir.T dan Ibu Eva Susanti, yang merupakan anak pertama dari dua bersaudara. Peneliti bertempat tinggal di Jalan Sejahtera, Kecamatan Pangkalan Kerinci, Kabupaten Pelalawan. Peneliti menempuh pendidikan dimulai dari SD Negeri 007 Pangkalan Kerinci pada tahun 2007 sampai 2013, selanjutnya peneliti meneruskan pendidikannya di SMP Negeri Bernas Kabupaten Pelalawan pada tahun 2013



Suska Riau
Sultan Syarif Kasim Riau

sampai 2016. Setamatnya Sekolah Menengah Pertama peneliti melanjutkan pendidikan di SMK Negeri 2 Pekanbaru dengan Jurusan Rekayasa Perangkat Lunak, dan peneliti melanjutkan pendidikan Strata Satu (S1) di Universitas Islam Negeri Sultan Syarif Kasim Riau pada Fakultas Sains dan Teknologi tepatnya pada Program Studi Sistem Informasi tahun 2019. Selama perkuliahan peneliti aktif dalam mengikuti berbagai seminar-seminar yang diadakan oleh kampus maupun di luar kampus. Selain itu, peneliti juga pernah menjadi anggota kepengurusan Himpunan Mahasiswa Sistem Informasi (HIMASI) tepatnya pada divisi Sumber Daya Manusia (SDM) dan juga menjadi kepanitian dalam acara Pendidikan Organisasi (PO) pada tahun 2021. Peneliti juga pernah mengikuti Kuliah Kerja Nyata di Desa Gabung Makmur, Kecamatan Kerinci Kanan, Kabupaten Siak pada tahun 2022. Pada penelitian Tugas Akhir ini peneliti mengambil topik Data Mining dengan judul penelitian Tugas Akhir "Komparasi Algoritma Data Mining Untuk *Clustering* Data Penderita Penyakit Menular Pada Manusia".

© Hak

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.