

DETEKSI DINI COVID-19 MELALUI CITRA CT-SCAN PARU-PARU MENGUNAKAN K-NEAREST NEIGHBOR DENGAN KOMPARASI JARAK

Lu'luul Maknun¹, Abdul Syukur², Affandy³ dan Moch Arief Soeleman⁴

Fakultas Ilmu Komputer, Universitas Dian Nuswantoro Semarang, Indonesia^{1,2,3 dan 4}

Email: luluulmaknunsuxes@gmail.com¹, abdulsyukur@gmail.com²,

affandy@gmail.com³ dan ariefsoeleman@gmail.com⁴

Abstrak

Covid -19 yang telah mewabah dan menjadi pandemik secara global yang merupakan masalah utama yang perlu di perhatikan dan di tangani, beberapa cara yang harus di lakukan adalah dengan memutus mata rantai penyebaran virus salah satunya dengan melakukan deteksi dini dan melakukan karantina, dengan CT scan paru-paru. CT scan paru-paru dapat dijadikan jalan alternatif. Berdasarkan permasalahan di atas maka peneliti mengetahui kondisi paru-paru secara detail dan dalam mendiagnosis virus secara dini. Pada penelitian ini pendekatan yang di ajukan menggunakan metode K-NN dengan perhitungan jarak *euclidean distance*, *manhattan distance*, *miskowski distance* untuk deteksi dini Covid -19 melalui citra CT scan paru-paru yang di duga terinfeksi Covid -19 . dalam mendeteksi secara dini evaluasi yang di gunakan untuk mengetahui pervorma yang di usulkan menggunakan *coufusion matrix* dengan hasil eksperimen menunjukkan hasil dari tiga perhitungan jarak menunjukkan hasil akurasi yang baik dan menggunakan dataset secara publik yaitu *euclidean distance* berjumlah 83%, *Manhattan distance* berjumlah 87%, Minkowski berjumlah 76%, di harapkan metode ini dapat di gunakan dan di kembangkan untuk melengkapi dioglosa medis.

Keywords: Covid-19; CT Scan; K-N; Euclidean Distance; Manhattan Distance; Miskowski Distance

Abstract

Covid -19 which has become an epidemic and has become a global pandemic which is the main problem that needs to be considered and handled, several ways that must be done are to break the chain of virus spread, one of which is by conducting early detection and quarantine, with a CT scan of the lungs. -lungs. CT scan of the lungs can be used as an alternative route. Based on the problems above, researchers know the condition of the lungs in detail and in diagnosing the virus early. In this study, the proposed approach uses the K-NN method with the calculation of the Euclidean distance, Manhattan distance, and Miskowski distance for early detection of Covid-19 through CT scan images of lungs suspected of being infected with Covid-19. in detecting early the evaluation used to determine the proposed performance using a *coufusion matrix* with experimental results showing the results of three distance calculations show good accuracy results and using a public dataset, namely the Euclidean distance of 83%, Manhattan distance of 87%, Minkowski amounting to 76%, it is hoped that this method can be used and developed to complete medical diagnosis.

Keywords: Covid-19; CT Scan; K-N; Euclidean Distance; Manhattan Distance; Miskowski Distance

Pendahuluan

Coronavirus Disease 2019 yang merupakan singkatan dari Covid-19 (Saputra, Arsyi, Nurhanifah, Octavia, & Pratomo, 2020). Jenis virus yang bersifat penyakit yang sangat menular seperti influenza yang disebabkan oleh *Severe Acute Syndrome Coronavirus 2* (SARS-Cov-2) (Haifa & MIA, 2020) dimana istilah Covid-19 pada tanggal 11 Februari 2020 telah di resmikan oleh *World Health Organization* (WHO) (Amin, Pinilih, & Astuti, 2021), gejala infeksi virus pada umumnya di kaitkan dengan infeksi saluran pernafasan (Fong, 2021) dan mempunyai gejala seperti demam, sakit kepala, sesak dan batuk (Abdillah, 2020) Covid -19 adalah jenis penyakit yang memiliki risiko penularan secara langsung ataupun tidak langsung (Hoesea, 2014) dan di sebabkan oleh jamur, parasit, bakteri, mikroorganisme (Syauqi, 2017) dan Covid-19 dapat di cegah dengan vaksinasi serta dengan pola hidup yang sehat (Suharmanto, 2020).

Menurut (Yanti & Hayatun, 2020) beberapa teknologi pencitraan medis, ada 3 metode dalam mendiagnosis Covid -19 di antara dengan melakukan tes darah, dengan sinar X, dan pemindahan computer yaitu (CT), CT merupakan teknologi yang non invasik, dan di anggap akurat dalam deteksi yang canggih, karakterisasi dan prositas internal. Berdasarkan permasalahan di atas maka harapan penelitian ini dapat membantua para medis dalam mendiagnosis pasien Covid-19 secara dini melalui gambar CT scan paru-paru.

Informasi mining merupakan ekstraksi data ataupun pola yang berarti ataupun menarik dari informasi yang terdapat di database yang besar (Eska, 2018). Informasi mining ialah bagian dari tahapan proses *Knowledge in Database* (KDD). Dengan informasi mining, kita bisa melaksanakan pengklasifikasian, memprediksi, memperkirakan serta memperoleh data lain yang berguna dari kumpulan informasi dalam jumlah yang besar (Mardi, 2017).

Pengelolaan citra merupakan cara dari berbagai modifikasi dan memanipulasi berbagai teknik. Pengelolaan citran sangatlah penting sebagai dasar di berbagai aplikasi yang nyata seperti penginderaan jarak jauh, pengenalan pola, dan machine vision dan pengelolaan sangat berperan dalam memiahkan latar belkangh dari objek secara otomatis (Kadir & Susanto, 2013).

Algoritma K-Nearest Neighbor (K-NN)

Classifier merupakan suatu tata cara buat melaksanakan klasifikasi terhadap objek bersumber pada informasi pendidikan yang jaraknya sangat dekat dengan objek tersebut. Informasi pendidikan diproyeksikan ke ruang berukuran banyak, yang setiap ukuran merepresentasikan fitur dari informasi. Ruang ukuran dipecah jadi bagian- bagian bersumber pada klasifikasi informasi pendidikan. Nilai k yang terbaik buat algoritma ini bergantung pada informasi, secara universal nilai k yang besar hendak kurangi dampak *noise* pada klasifikasi, hendak namun membuat batas antara tiap klasifikasi jadi lebih buram. Nilai k yang bagus bisa diseleksi dengan optimis parameter, misalnya dengan memakai *cross-validation*. Permasalahan spesial buat klasifikasi diprediksikan bersumber pada informasi pendidikan yang sangat dekat (dengan kata lain, $k=1$) yang umumnya diucap algoritma *nearest neighbor* (Wijaya, Irsyad, & Widhiarso, 2020) Tata cara Euclidean Distance ialah penghitungan jarak pada algoritma KNN yang sangat banyak digunakan oleh peneliti.

Rumus Euclidean Distance:

$$n = \sqrt{\sum_{i=1} (x_i - y_i)^2}$$

Dengan

- a) $d(x,y)$ = jarak antara x dan y
- b) x_i = *Data testing* ke-i
- c) y_i = *data training* ke- i
- d) n = dimensi data

Ct-Scant

CT- scan paru- paru akan bisa menolong penegakan penaksiran Covid-19 dan juga bagikan cerminan CT-scan pada pasien- pengidap Covid-19 (Restuningdyah & Amalia, 2020)

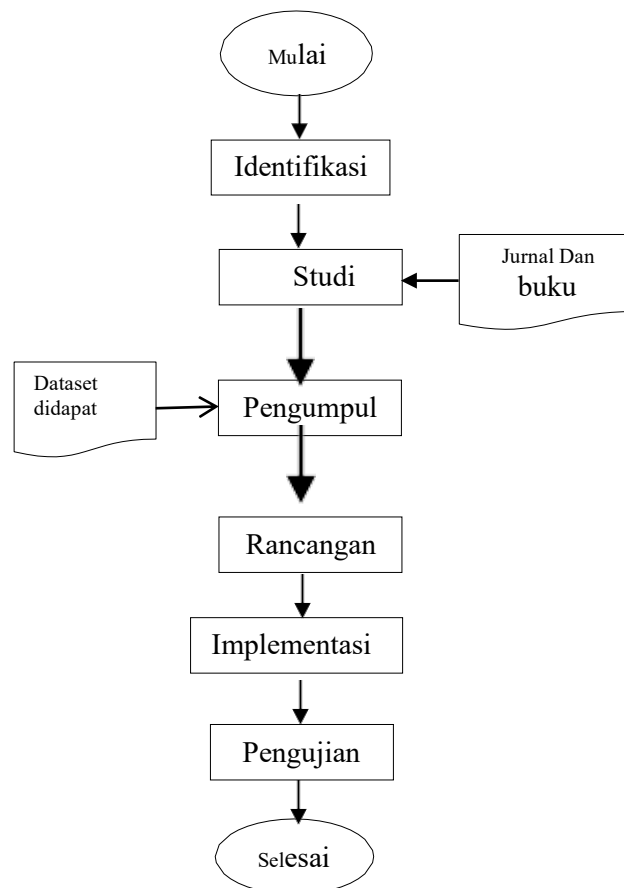
Komparasi Jarak

Koordinat yang universal digunakan dalam membetulkan posisi dalam peta yakni latitude serta longitude yakni garis yang melintang Yusup Miftahuddin, dkk Tekno Insentif- 71 diantara

kutub utara serta selatan yang menghubungkan sisi timur serta barat bagian bumi (khatulistiwa). Sebaliknya *Longitude* yaitu garis yang menghubungkan antara sisi utara serta sisi selatan bumi (kutub) dalam matematika *euclidean distance* digunakan buat mengukur 2 titik dalam satu ukuran yang membagikan hasil semacam perhitungan Pythagoras (Miftahuddin, Umaroh, & Karim, 2020). *Minkowski distance* adalah metrik dalam suatu vector yang di anggapa sebagai generasi dari *Euclidean distance* dan *Manhattan distance*. *Manhattan distance* ialah perhitungan yang membedakan secara absolut (mutlak) antara sepasang kordinat objek. *Euclidean distance* ialah perhitungan buat mengukur 2 jarak titik dari *eunclidean* ruang di antara 2 sudut

Metode Penelitian

Desain sistem yang akan dibangun di dalam penyusunan tesis ini. Desain sistem secara umum mengenai kontribusi penelitian deteksi fraud dapat dilihat pada:



Gambar 1. Tahapan Penelitian.

Data citra scan paru-paru.

Refrensi dari jurnal internasional yang sudah melakukan penelitian sebelumnya pada pasien melalui CT Scan paru-paru 50 Covid-19, 50 non Covid-19.

Deteksi Dini Covid-19 Melalui Citra CT-Scan Paru-Paru Menggunakan K-Nearest Neighbor dengan Komparasi Jarak



Gambar 2. Data Citra Scan Paru-Paru.

Hasil dan Pembahasan

Berdasarkan perhitungan jarak *euclidean distance*, *manhattan distance*, *miskowski distance* yang di gunakan untuk menghitung jarak terdekat antara data citra covid-19 positif dan data citra covid-19 negatif terlihat pada table 4.4 dengan menghitung persamaan sehingga memperoleh hasil sebagai berikut

Data Latih	Hasil Perhitungan Jarak Euclidean, Manhattan, Miskowski		
Covid -19	8.660254037844387	0.0	0.0
Covid -19	1.7320508075688772	3.0	3.0
Covid -19	10.392304845413264	15.0	15.0
Covid -19	20.784609690826528	18.0	18.0
Covid -19	8.660254037844387	36.0	36.0

Gambar 3. Jarak Data Uji dengan Citra Covid-19.

Data Latih	Hasil Perhitungan Jarak Euclidean, Manhattan, Miskowski		
Covid -19	8.660254037844387	0.0	0.0
Covid -19	1.7320508075688772	3.0	3.0
Covid -19	10.392304845413264	15.0	15.0
Covid -19	20.784609690826528	18.0	18.0
Covid -19	8.660254037844387	36.0	36.0

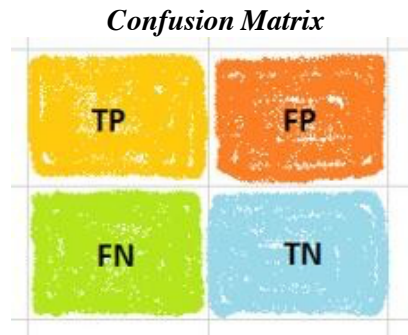
Gambar 4. Jarak Data Uji dengan Citra Non Covid-19.

Metode yang akan di gunakan pada penentuan klasifikasi kelas dari citra CT Scan paru-paru bertujuan untuk mengetahui jenis kelasnya, dalam penentuan kelas ada dua jenis kelas klasifikasi covid 19 dan non covid 19, dan hasil dari euclidean distance, manhattan distance, minkowski distance pada table 4.3 dengan 4.4, yang di gabung kemudian dalam menentukan nilai tetangga terdekat, dalam rangking di nyatakan tingkat kesamaanya dalam penelitian ini k=3

menjadi ranking tiga terdekat atau terkecil jadi semakin dengan kesamaan yang semakin dekat, maka jumlah jumlah rata-rata menentukan jenis kelasnya, hasil ranking perhitungan jarak *euclidean distance*, *manhattan distance*, *minkowski distance* pada gambar 3, gambar 4.

Pada sub bab analisa yang di dihasilkan ini telah menjelaskan tentang hasil dari metode yang telah diusulkan dan hasil dari analisa akan diperlihatkan secara tertulis dari seluruh hasil dari proses implementasi uji coba beserta analisisnya. Pada penelitian ini menggunakan *Confusion Matrix* untuk mengevaluasi sisten, bentuk dari *Confusion Matrix* di mana yang berupa *table matrix* dengan menggunakan kinerja model klasifikasi dengan kombinasi 4 nilai prediksi yaitu *True Positif (TP)*, *False Positif (FP)*, *True Negatif (TN)*, *False Negatif (FN)*

Dengan *table Confusion Matrix* dengan menghasilkan nilai *Accuracy*, *precision*, dan *Recall*



Gambar 5. *Confusion Matrix* Kalsifikasi.

Hasil dari klasifikasi pada data uji *Confusion Matrix* yang telah di peroleh yaitu *False Negatif (FN)* dengan nilai 0, *False Positif (FT)* dengan nilai 1, *True Negatif* dengan nilai 2, *True Positif (TP)* dengan nilai 2. Berdasarkan klasifikasi *Confusion Matrix* K-NN di tunjukkan pada Gambar 7.

2	1
0	2

Total	Jarak Euclidean	Jarak Manhattan	Jarak Minkowski
True Positive (TP)	18	18	17
False Positive (FP)	5	2	5
False Negative (FN)	2	3	5
True Negative (TN)	15	15	15

Gambar 7. *Confusion Matrix* Klasifikasi K-NN.

Akurasi = $(TP + TN) / (TP + TN + FP + FN)$
Presisi = $TP / (TP + FP)$
Sensitifitas = $TP / (TP + FN)$
Specificity = $TN / (TN + FP)$

Hasil	Jarak Euclidean	Jarak Manhattan	Jarak Minkowski
Akurasi	83%	87%	76%
Presisi	78%	90%	77%
Sensitifitas	90%	86%	77%
Specificity	75%	88%	75%

Gambar 8. Evaluasi Kerja Sistem.

Bibliografi

- Abdillah, Leon. (2020). Stigma Terhadap Orang Positif COVID-19 (Stigma on Positive People COVID-19). *Pandemik COVID-19: Antara Persoalan Dan Refleksi Di Indonesia, Forthcoming*.
- Amin, Muhammad Khoirul, Pinilih, Sambodo Sriadi, & Astuti, Retna Tri. (2021). Gambaran Psikologi Warga Kabupaten Magelang Selama Pandemi Covid-19. *Jurnal Mutiara Ners, 4*(2), 140–145.
- Eska, Juna. (2018). *Penerapan Data Mining Untuk Prediksi Penjualan Wallpaper Menggunakan Algoritma C4. 5*.
- Fong, Karen Lee Mei. (2021). *Antimicrobial sensitivity of most commonly isolated bacteria from Feline Upper Respiratory Infection (URI)*.
- Haifa, Najwa S., & MIA, X. (2020). Pendidikan Kesehatan Tentang Penyakit Menular. *OSF Preprints, May, 9*.
- Hoesea, Elvan Virgo. (2014). Evaluation of health surveillance activities of hajj 2013 in the hajj embarkation Palangkaraya. *Jurnal Berkala Epidemiologi, 2*(2), 206–215.
- Kadir, Abdul, & Susanto, Adhi. (2013). Teori dan aplikasi pengolahan citra. *Yogyakarta: Andi*.
- Mardi, Yuli. (2017). Data Mining: Klasifikasi Menggunakan Algoritma C4. 5. *Jurnal Edik Informatika Penelitian Bidang Komputer Sains Dan Pendidikan Informatika, 2*(2), 213–219.
- Miftahuddin, Yusup, Umaroh, Sofia, & Karim, Fahmi Rabiul. (2020). Perbandingan Metode Perhitungan Jarak Euclidean, Haversine, Dan Manhattan Dalam Penentuan Posisi Karyawan. *Jurnal Tekno Insentif, 14*(2), 69–77.
- Restuningdyah, Novia Andansari Putri, & Amalia, Emmy. (2020). Pemeriksaan CT-SCAN Thorax Pada Kasus Covid-19 di Provinsi Nusa Tenggara Barat. *Jurnal Pengabdian Magister Pendidikan IPA, 3*(1).
- Saputra, Maman, Arsyi, Miftahul, Nurhanifah, Nurhanifah, Octavia, Syally Nadya, & Pratomo, Hadi. (2020). Evaluasi pedoman penanganan cepat medis dan kesehatan masyarakat tentang coronavirus disease (covid-19) di indonesia. *Jurnal Ilmiah Ilmu Keperawatan Indonesia, 10*(02), 46–55.
- Suharmanto, Suharmanto. (2020). Perilaku Masyarakat dalam Pencegahan Penularan Covid-19. *JK Unila, 4*(2), 91–96.
- Syauqi, Ahmad. (2017). *Mikrobiologi lingkungan peranan mikroorganisme dan kehidupan*. Penerbit Andi.
- Wijaya, Chandra, Irsyad, Hafiz, & Widhiarso, Wijang. (2020). Klasifikasi Pneumonia Menggunakan Metode K-Nearest Neighbor Dengan Ekstraksi GLCM. *Jurnal Algoritme, 1*(1), 33–44.
- Yanti, Budi, & Hayatun, Ulfa. (2020). Peran pemeriksaan radiologis pada diagnosis Coronavirus disease 2019. *Jurnal Kedokteran Syiah Kuala, 20*(1).