

KLASIFIKASI TINGKAT LAJU DATA COVID-19 UNTUK MITIGASI PENYEBARAN MENGGUNAKAN METODE *MODIFIED K-NEAREST NEIGHBOR* (MKNN)

Imam Cholissodin^{*1}, Felicia Marvela Evanita², Jeffrey Junior Tedjasulaksana³, Kukuh Wicaksono Wahyuditomo⁴

¹²³⁴Fakultas Ilmu Komputer, Universitas Brawijaya

Email: ¹imamcs@ub.ac.id, ²fmarvela@student.ub.ac.id, ³jeffrey@student.ub.ac.id, ⁴kukuhw@student.ub.ac.id

*Penulis Korespondensi

(Naskah masuk: 23 November 2020, diterima untuk diterbitkan: 09 Juni 2021)

Abstrak

COVID-19 atau *Coronavirus Disease 2019* merupakan sebuah penyakit yang disebabkan oleh virus yang dapat menular melalui saluran pernapasan pada hewan atau manusia dan menyebabkan ribuan orang meninggal hampir di seluruh dunia, sehingga dinyatakan sebagai sebuah pandemi di banyak negara, termasuk di Indonesia. Kasus COVID-19 pertama kali ditemukan di Indonesia pada tanggal 2 Maret 2020, dalam menangani pandemi COVID-19 pemerintah menerapkan *social distancing* dengan menjaga jarak antara satu sama lain sejauh lebih dari 1 meter dan menerapkan protokol kesehatan yang telah diatur saat melakukan aktivitas di luar rumah sesuai anjuran *World Health Organization* (WHO). Rendahnya kesadaran masyarakat Indonesia dalam menerapkan *social distancing* dan protokol kesehatan menyebabkan bertambahnya kasus positif COVID-19 di Indonesia secara signifikan sehingga banyak korban yang meninggal, oleh karena itu pada penelitian ini kami membuat sistem klasifikasi tingkat laju data COVID-19 untuk mitigasi penyebaran di seluruh provinsi di Indonesia dengan menggunakan metode *Modified K-Nearest Neighbor* (MKNN) dengan hasil keluaran berupa kelas laju penyebaran yaitu laju penyebaran rendah yang artinya mitigasi penyebarannya tinggi, kemudian kelas laju penyebaran sedang yang artinya mitigasi penyebarannya sedang, dan laju penyebaran tinggi yang berarti mitigasi penyebaran rendah dan dijelaskan lebih lanjut pada bagian metodologi penelitian. Hasil keluaran dari sistem bertujuan untuk meningkatkan kesadaran masyarakat Indonesia dalam mencegah COVID-19 dengan melihat kelas laju penyebaran pada masing-masing provinsi di Indonesia. Alasan penggunaan metode *Modified K-Nearest Neighbor* pada penelitian ini adalah karena metode *Modified K-Nearest Neighbor* merupakan salah satu metode klasifikasi yang cukup baik, dimana pada metode ini dilakukan pemvalidasian dan pembobotan yang bobotnya ditentukan dengan menghitung fraksi dari tetangga berlabel yang sama dengan total jumlah tetangga. Parameter yang digunakan dalam proses klasifikasi adalah jumlah kasus positif, jumlah orang yang sembuh, dan jumlah orang yang meninggal akibat COVID-19. Data yang digunakan pada penelitian ini berasal dari situs resmi kementerian kesehatan republik Indonesia yang dapat diakses pada link <https://infeksiemerging.kemkes.go.id/> dengan jumlah data latih sebanyak 374 data pada tanggal 12 Mei 2020 sampai 22 Mei 2020 dan data uji sebanyak 136 data pada tanggal 23 Mei 2020 sampai tanggal 26 Mei 2020, hasil akurasi yang dihasilkan adalah 97,79% dengan nilai $K = 3$.

Kata kunci: klasifikasi, COVID-19, modified k-nearest neighbor

THE CLASSIFICATION OF COVID-19 DATA RATE FOR DISTRIBUTION MITIGATION USING MODIFIED K-NEAREST NEIGHBOR (MKNN)

Abstract

COVID-19 or *Coronavirus 2019* is a disease caused by a virus that can be transmitted through the respiratory tract to animals or humans and causes more people to die around the world, making it a pandemic in many countries, including Indonesia. COVID-19 cases were first discovered in Indonesia on March 2, 2020. Under the COVID-19 pandemic agreement, the government imposed a social grouping with a grouping of more than 1 meter apart from one another and the transfer of related health protection when carrying out activities outside the home as directed by the *World Health Organization* (WHO). Considering the Indonesian people in implementing social preservation and protecting health policies increase the positive acquisition of COVID-19 in Indonesia significantly related to the number of victims who died, therefore in this study, we created a COVID-19 data level assessment system for transfer mitigation in all provinces in Indonesia by using the *Modified K-Nearest Neighbor* (MKNN) method with the output in the form of a spread rate class, namely a low

spread rate which means that the spread mitigation is high, then the medium spread rate class which means the spread mitigation is moderate, and the spread rate is high which means low spread mitigation which is further explained in the section on the research methodology. The purpose of the system output is to increase the awareness of the Indonesian people in preventing COVID-19. The parameters used in the classification process are the number of positives, the number of people recovered, and the number of people died by COVID-19 by looking at the class distribution rate in each province in Indonesia. The reason for using the Modified K-Nearest Neighbor method in this research is because the Modified K-Nearest Neighbor method is a fairly good classification method, where this method is validated and weighted whose weight is determined by calculating the fraction of neighbors labeled the same as the total of neighbors number. The data used in this study was released from the official website of the Ministry of Health of the Republic of Indonesia which can be accessed at the link <https://infection.infemerging.kemkes.go.id/> with a total of 374 training data from May 12, 2020 to May 22, 2020 and test data As many as 136 data from 23 May 2020 to 26 May 2020, the resulting accuracy was 97.79% with a $K = 3$.

Keywords: *classification, COVID-19, modified k-nearest neighbor*

1. PENDAHULUAN

COVID-19 merupakan jenis coronavirus yang pertama kali ditemukan di Wuhan, Tiongkok, bulan Desember 2019 dan dapat menular melalui saluran pernapasan pada hewan atau manusia, COVID-19 juga dinyatakan sebagai sebuah pandemi di banyak negara di seluruh dunia termasuk Indonesia (Setiati, 2020).

Kasus COVID-19 di Indonesia pertama kali ditemukan pada tanggal 2 Maret 2020 kemudian menyebar sehingga menyebabkan ribuan orang meninggal dan terus bertambah tiap harinya. Oleh karena itu, sesuai anjuran *World Health Organization* (WHO) pemerintah Indonesia menerapkan *social distancing* dimana tiap warga Indonesia harus menjaga jarak sejauh 2 meter dan mematuhi protokol kesehatan seperti memakai masker, rajin mencuci tangan, melakukan deteksi suhu tubuh saat berada di tempat umum agar terhindar dari COVID-19. Akan tetapi masih banyak masyarakat Indonesia yang masih tidak melakukan *social distancing* atau mematuhi protokol kesehatan saat berada di tempat umum seperti di pasar atau saat menggunakan transportasi umum karena rendahnya kesadaran dan kewaspadaan masyarakat Indonesia. Hal ini terbukti dari jumlah angka kematian yang dipaparkan dari web resmi Kementerian Kesehatan Indonesia per tanggal 20 Mei 2020 mencapai 1.242 jiwa sejak kasus COVID-19 pertama di Indonesia.

Berdasarkan masalah tersebut pada penelitian ini dilakukan klasifikasi tingkat laju data COVID-19 pada tiap provinsi untuk meningkatkan kesadaran masyarakat Indonesia menggunakan metode MKNN. Metode MKNN merupakan salah satu metode yang cukup baik dalam mengklasifikasikan data. Penelitian dalam melakukan klasifikasi dengan menggunakan metode MKNN telah digunakan pada penelitian Identifikasi Jenis *Attention Deficit Hyperactivity Disorder* (ADHD) pada anak usia dini (Ariyanti, et al., 2019) menghasilkan akurasi terbaik sebesar 90% saat menggunakan data latih sebanyak 80 data dan data uji sebanyak 20 data dengan nilai $K = 3$.

Penelitian selanjutnya menggunakan MKNN untuk Klasifikasi Status Gunung Berapi (Anggian, et al., 2019). Pada penelitian tersebut menggunakan 110 data yang diperoleh dari badan resmi yang berwenang dalam mengeluarkan status gunung berapi yaitu Pusat Vulkanologi dan Mitigasi Geologi (PVMBG). Hasil pengujian terbaik dari penelitian tersebut adalah 86,67% dan untuk rata-rata akurasi sebesar 82,87% dengan menggunakan beberapa nilai K yaitu 3, 5, 6, 7, dan 9.

Selanjutnya penelitian terkait desain *smart electrocardiography* (ECG) menggunakan MKNN (Yuwono, et al., 2018). Pada desain *smart electrocardiography* (ECG) berfungsi sebagai mendiagnosis penyakit jantung berdasarkan interpretasi sinyal ECG, jenis jantung dibagi menjadi 3 kelas, yaitu normal, infark miokard, dan lainnya. Data ECG yang digunakan pada penelitian ini bersumber dari MIT BIH dan menghasilkan akurasi terbaik pada saat $K = 3$ sebesar 71,2%.

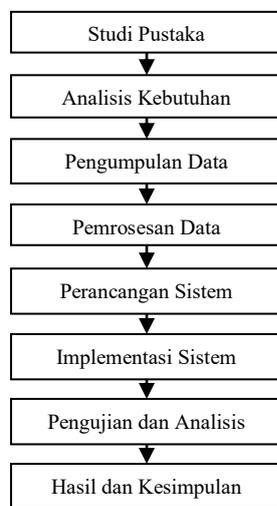
Penelitian MKNN selanjutnya adalah klasifikasi gen kanker (Ayyad, et al., 2019). Penelitian ini menggunakan data gen yang terdiri dari puluhan sampel yang telah ditandai oleh ribuan gen dan dilakukan percobaan pada enam dataset ekspresi gen yang berbeda dan lebih unggul dibandingkan dengan metode KNN, *weighted KNN*, SVM, SVM Fuzzy, dan brain emotional learning dalam hal klasifikasi akurasi, presisi, dan *recall*. Metode MKNN memiliki waktu pengujian yang lebih kecil dibandingkan KNN dan *weighted KNN*.

Berdasarkan pada beberapa penelitian sebelumnya terkait MKNN dan permasalahan tentang COVID-19, maka pada penelitian ini kami membuat program dalam mengklasifikasikan tingkat laju data COVID-19 dalam mitigasi penyebaran pada tiap provinsi di Indonesia dengan menggunakan metode MKNN untuk meningkatkan kesadaran dan kewaspadaan masyarakat Indonesia dalam menghadapi pandemi COVID-19 dengan menggunakan beberapa parameter yaitu jumlah kasus positif, jumlah pasien yang sembuh, dan jumlah pasien yang telah meninggal. Pada penelitian ini menggunakan data sebanyak 374 data pada

tanggal 12 sampai 26 Mei 2020 yang diperoleh dari Kementerian Kesehatan.

2. METODE PENELITIAN

Dalam penelitian ini dilakukan klasifikasi pada hasil rekap data perkembangan kasus COVID-19 di Indonesia. Klasifikasi adalah proses mengolah data untuk menemukan sebuah model dengan tujuan untuk memperkirakan kelas dari objek yang tidak diketahui kelasnya (Parapat, 2018). Metode klasifikasi yang digunakan pada penelitian ini adalah metode *Modified K-Nearest Neighbor*. Metode *Modified K-Nearest Neighbor* adalah metode yang dikembangkan dari metode *K-Nearest Neighbor* Pada Gambar 1 ditunjukkan alur dari metode penelitian ini.



Gambar 1. Metode Penelitian

2.1 Dataset

Dataset yang digunakan pada penelitian ini adalah rekap data perkembangan kasus COVID-19 setiap provinsi yang ada di Indonesia, dimulai dari tanggal 12 Mei 2020 hingga 26 Mei 2020, yang di mana diambil dari situs web resmi Kementerian Kesehatan Republik Indonesia yang dapat diakses dari <https://infeksiemerging.kemkes.go.id>. Data yang digunakan merupakan kumpulan data dengan komposisi fitur sejumlah 3, yaitu positif, sembuh, dan meninggal. Dari 510 data yang dikumpulkan, 70% porsi data dijadikan sebagai data latih dan 30% porsi data dijadikan sebagai data uji, sehingga jumlah data latih yang digunakan adalah sebanyak 374 data dan jumlah data uji yang akan digunakan adalah sebanyak 136 data. Berikut Tabel 1 yang merupakan sampel data dari sebuah provinsi pada *dataset* yang digunakan oleh penelitian ini.

2.2 Modified K-Nearest Neighbor

Modified K-Nearest Neighbor (MKNN) merupakan metode klasifikasi pengembangan dari metode *K-Nearest Neighbor* (KNN) dengan menggunakan tetangga yang kuat dalam data pelatihan. Penggunaan metode MKNN tidak berbeda jauh dengan metode KNN standar seperti yang dilakukan oleh Marji (2019) untuk klasifikasi penyakit kanker (Marji, et al., 2019), di mana dilakukan pengklasifikasian tes sample sesuai dengan kemiripan fitur tetangganya, tetapi pada metode MKNN dilakukan pemvalidasian dan pembobotan yang bobot nya ditentukan dengan menghitung fraksi dari tetangga berlabel yang sama dengan total jumlah tetangga. Hasil akurasi pada metode MKNN lebih tinggi dibandingkan metode KNN (Parvin, et al., 2010). Berikut proses metode MKNN (Yuwono, 2018):

1. Menentukan nilai K yang ingin digunakan pada pengujian (H tetangga). Nilai K yang biasa digunakan merupakan angka ganjil.
2. Melakukan perhitungan menggunakan euclidean distance pada data latih untuk mengetahui jarak euclidean antar data latih yang digunakan, seperti yang ditunjukkan pada Persamaan 1.

$$d(x_i, x_j) = \left(\sum_{l=1}^{nDim} |x_{il} - x_{jl}|^2 \right)^{\frac{1}{2}} \quad (1)$$

Keterangan:

x_i = data latih ke- i

x_j = data latih ke- j

i, j = indeks data latih

l = indeks fitur data

$nDim$ = banyaknya fitur data

x_{il} = data latih ke- i pada fitur ke- l

x_{jl} = data latih ke- j pada fitur ke- l

3. Melakukan perhitungan untuk menentukan nilai validitas pada data latih, di mana nilai K yang digunakan dihitung dari data latih dengan jarak terdekat. Rumus menghitung nilai validitas data latih dapat dilihat pada Persamaan 2.

$$Validity(x) = \frac{1}{H} \sum_{i=1}^H S(lbl(x), lbl(N_i(x))) \quad (2)$$

Keterangan:

x = data latih

H = jumlah titik terdekat atau nilai K

$lbl(x)$ = label kelas aktual dari data x

$lbl(N_i(x))$ = label kelas titik terdekat x ketika tetangga terdekatnya sebanyak i

Tabel 1. Dataset Penelitian

NO	DAERAH	TANGGAL	POSITIF	SEMBUH	MENINGGAL	LABEL
1	ACEH	12 May	0	1	0	Rendah
2		13 May	0	0	0	Rendah
3		14 May	0	1	0	Rendah
4		15 May	0	0	0	Rendah
5		16 May	1	2	0	Rendah
6		17 May	0	0	0	Rendah
7		18 May	0	0	0	Rendah
8		19 May	0	0	0	Rendah
9		20 May	0	0	0	Rendah
10		21 May	0	0	0	Rendah

Rumus menentukan kesamaan antar titik data pada tetangga terdekat dapat dilihat pada Persamaan 3.

$$S(a, b) = \begin{cases} 1 & a = b \\ 0 & a \neq b \end{cases} \quad (3)$$

Keterangan:

$$a = lbl(x)$$

$$b = lbl(N_i(x))$$

- Melakukan perhitungan menggunakan euclidean distance pada data uji untuk mengetahui jarak antar data uji dengan data latih.
- Melakukan perhitungan weight voting dari K tetangga terdekat dan melakukan validitas pada data latih yang akan dilakukan perhitungan perkalian dengan weight voting pada setiap tetangga. Berikut Persamaan 4 merupakan rumus weight voting atau nilai bobot.

$$W(c_i) = Validity(c_i) \times \frac{1}{d_e + \alpha} \quad (4)$$

Keterangan:

$W(c_i)$ & $Validity(c_i)$ = nilai bobot & validitas dari data latih c_i yang merupakan salah satu tetangga terdekat dari data uji, di mana $i = 1, 2, \dots, K$

d_e = jarak euclidean antara data latih dengan data uji

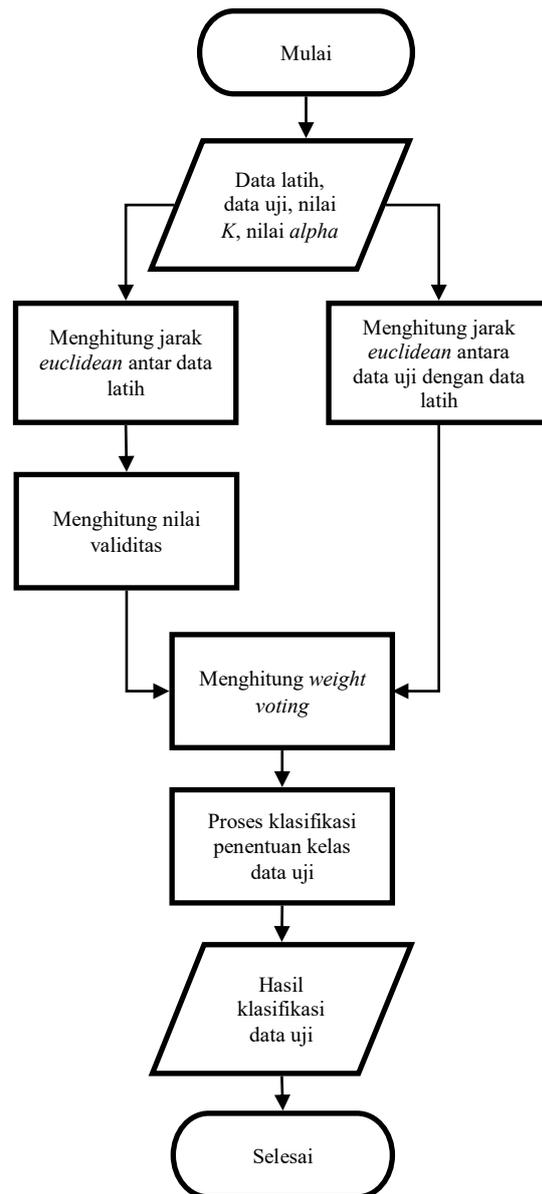
Alpha (α) = nilai regulator smoothing (perataan), secara umum nilai yang digunakan sebesar 0,5

- Dan langkah terakhir adalah menentukan kelas dari data uji berdasarkan hasil perhitungan weight voting dengan melihat nilai weight voting terbesar dan diurutkan secara *descending* sesuai dengan nilai K yang telah ditentukan sebelumnya.

Proses dari metode *Modified K-Nearest Neighbor* secara keseluruhan dalam bentuk *flowchart*, ditunjukkan melalui Gambar 2 berikut ini.

3. HASIL DAN PEMBAHASAN

Dari penelitian yang dikerjakan, hasil yang diperoleh akan melalui pengujian, dimana melalui beberapa parameter yaitu melalui pengujian *confusion matrix*, pengujian pengaruh nilai K , dan pengaruh pengujian pengaruh nilai $alpha$.



Gambar 2. Flowchart Metode Modified K-Nearest Neighbor

3.1 Implementasi Sistem

Implementasi dari penggunaan metode *Modified K-Nearest Neighbor* dalam penelitian ini, memungkinkan untuk memberi masukan data baru terkait laju data dari COVID-19 untuk diklasifikasikan tingkatnya. Masukan yang dibutuhkan mengikuti fitur yang digunakan, yaitu

jumlah pasien yang positif, jumlah pasien yang sembuh, dan jumlah pasien yang meninggal dunia. Selain data baru, dapat juga menentukan nilai K dan nilai α sebagai bagian dalam proses perhitungan. Implementasinya dapat dilihat dari Gambar 3.



Gambar 3. Hasil Implementasi *Modified K-Nearest Neighbor*

3.2. Pengujian *Confusion Matrix*

Pada pengujian berikut melibatkan nilai K sebesar 3 dengan α sebesar 0,5. Karena hasil proses klasifikasi melibatkan kelas lebih dari dua, maka digunakan model *multiclass confusion matrix* untuk mengumpulkan hasil klasifikasinya, seperti yang ditampilkan pada Tabel 2 di bawah ini.

Tabel 2. Hasil *Confusion Matrix*

		Aktual		
		Rendah	Sedang	Tinggi
Hasil Pengujian	Rendah	125	2	0
	Sedang	0	8	1
	Tinggi	0	0	0

Hasil Tabel 2 diolah kembali melalui 4 parameter seperti yang ditunjukkan pada Tabel 3 sebagai bagian untuk melihat kinerja metode *Modified K-Nearest Neighbor* sebagai sebuah metode klasifikasi data, yaitu *precision*, *recall*, *F-measure*, dan *specificity*, dengan hasil dari label kelas laju data rendah memberikan kinerja yang cukup baik dengan 3 parameter lebih besar dari kelas data yang lain, yaitu parameter *precision*, *recall*, dan *F-Measure*. Hal ini dipengaruhi dari banyaknya data yang memiliki kelas laju data penyebaran virus COVID-19 rendah yang berhasil diklasifikasikan oleh sistem seluruhnya, dibanding kelas data lain yang mengalami kesalahan klasifikasi.

Tabel 3. Hasil *Precision*, *Recall*, *F-Measure*, dan *Specificity*

Kelas	Precision	Recall	F-Measure	Specificity
Rendah	0,984251 9685	1	0,9920634 921	0,8
Sedang	0,888888 8889	0,8	0,8421052 632	0,9920634 921
Tinggi	0	0	0	1

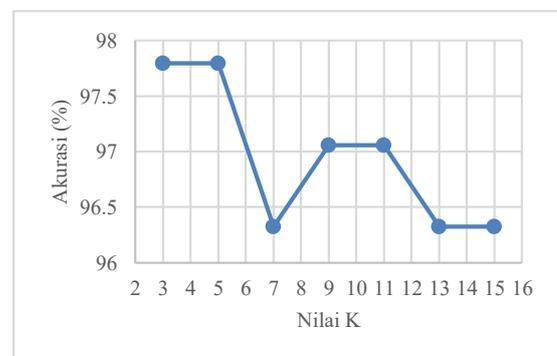
3.3 Pengujian Pengaruh Nilai K

Berdasarkan pengaruh jumlah data yang cukup berdekatan dengan data uji yang digunakan dalam pengujian, membuat pengaruh nilai K menghasilkan perbedaan terhadap nilai akurasi hasil pengujian proses klasifikasi yang diberikan oleh sistem, seperti yang tercantum pada Tabel 4.

Tabel 4. Hasil Pengujian Terhadap Pengaruh Nilai K

Nilai K	Akurasi Pengujian (%)
3	97,7941176470588
5	97,7941176470588
7	96,3235294117647
9	97,0588235294117
11	97,0588235294117
13	96,3235294117647
15	96,3235294117647

Hasil pengujian membuktikan, semakin besar nilai K yang digunakan, semakin kecil juga akurasi yang diberikan, seperti yang ditunjukkan Gambar 4.



Gambar 4. Grafik Hasil Pengujian Terhadap Pengaruh Nilai K

Penurunan nilai akurasi yang terjadi, juga dipengaruhi oleh tidak seimbang data dengan label kelas seperti tinggi dan sedang. Seperti pada saat nilai K yang digunakan semakin besar, ada data dari dengan kelas tersebut yang ikut masuk, sehingga membuat proses klasifikasi menjadi kurang akurat keputusannya.

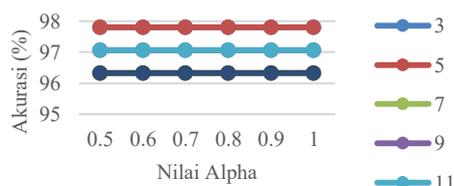
3.4 Pengujian Pengaruh Nilai α

Nilai pengujian pada α dimulai dari 0,5 hingga 1 yang melibatkan perbedaan pada setiap nilai K -nya seperti dari pengujian yang dilakukan sebelumnya, memberikan persentase hasil nilai akurasi yang ditampilkan dan dapat dilihat keseluruhannya pada Tabel 5.

Tabel 5. Hasil Pengujian Terhadap Pengaruh Nilai α

K	Akurasi Pengujian Dengan α					
	0.5	0.6	0.7	0.8	0.9	1
3	97,7941176	97,7941176	97,7941176	97,7941176	97,7941176	97,7941176
5	97,7941176	97,7941176	97,7941176	97,7941176	97,7941176	97,7941176
7	96,3235294	96,3235294	96,3235294	96,3235294	96,3235294	96,3235294
9	97,0588235	97,0588235	97,0588235	97,0588235	97,0588235	97,0588235
11	97,0588235	97,0588235	97,0588235	97,0588235	97,0588235	97,0588235
13	96,3235294	96,3235294	96,3235294	96,3235294	96,3235294	96,3235294
15	96,3235294	96,3235294	96,3235294	96,3235294	96,3235294	96,3235294

Tidak terdapat perbedaan nilai akurasi dari keseluruhannya, membuktikan perubahan terhadap nilai α sebagai *regulator smoothing* tidak berpengaruh besar terhadap akurasi dari proses klasifikasi. Karena pengaruhnya yang tidak terlalu besar, maka setiap proses perhitungan akan memberikan efek yang sama ketika ada keterlibatan nilai α , terlebih nilai α yang digunakan dalam proses umumnya memang kecil, sehingga tidak terlalu signifikan terhadap variabel lainnya. Secara ilustrasi, grafik pengaruh dari nilai α hasilnya dapat terlihat dari Gambar 5 berikut.

Gambar 5. Grafik Pengujian Terhadap Pengaruh Nilai α

4. KESIMPULAN DAN SARAN

Berdasarkan dari implementasi dan pengujian, diperoleh beberapa poin kesimpulan, yaitu nilai akurasi maksimal dari sistem yang diperoleh setelah pengujian dengan nilai K sebanyak 3 dan α 0,5 adalah 97,794%. dan pengujian paling maksimal ditunjukkan kelas laju data rendah dengan *precision* sebesar 0,984, *recall* sebesar 1, *F-Measure* sebesar 0,992, dan *specificity* sebesar 0,8, serta pengujian terhadap nilai K , diperoleh akurasi optimal ketika nilai K yang digunakan berjumlah 3 dan 5, yaitu sebesar 97,794%. Kemudian pengujian terhadap nilai α , tidak berpengaruh ke dalam hasil akurasi yang diberikan sistem. Sebagai saran, sebaiknya dilakukan penambahan data latih baru yang memiliki laju penyebaran tinggi, atau melibatkan pelabelan data ulang, dan dilakukan optimasi terhadap pemilihan nilai K atau terhadap nilai α sebagai *smoothing regulator*.

DAFTAR PUSTAKA

SETIATI, S. dan AZWAR, M.K., 2020. COVID-19 and Indonesia. *Acta Medica Indonesiana*, 52(1), pp.84-89.

WHO, 2019. Coronavirus. [online] Tersedia di:

<<https://www.who.int/indonesia/news/novel-coronavirus/qa-for-public>> [Diakses 03 June 2009].

ARIYANTI, R.N. dan INDRIATI, R.C.W., 2019. Identifikasi Jenis Attention Deficit Hyperactivity Disorder (ADHD) Pada Anak Usia Dini Menggunakan Metode Modified K-Nearest Neighbor (MKNN). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer and e-ISSN*, 2548, p.964X

ANGGIAN, F.C., HIDAYAT, N. FURQON, M.T., 2019. Implementasi Metode Modified K-Nearest Neighbor untuk Klasifikasi Status Gunung Berapi. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN*, 2548, p.964X.

MARJI, M., CHOLISSODIN, I., RATNAWATI, D. E., SANTOSO, E., and HIDAYAT, N., 2019. Features Reduction Using Deep Miden Algorithm for Classification of Cancer Based on Itemset Sequent of Protein TP53. *International Conference on Sustainable Information Engineering and Technology (SIET)*, Lombok, Indonesia, 2019, pp. 349-352, doi: 10.1109/SIET48054.2019.8986015.

PARAPAT, I. M., FURQON, M.T., SUTRISNO, 2018. Penerapan Metode Support Vector Machine (SVM) pada Klasifikasi Penyimpangan Tumbuh Kembang Anak. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN*, 2548, 964X.

PARVIN, H., ALIZADEH, H. and MINAEI-BIDGOLI, B., 2010, October. MKNN: Modified k-nearest neighbor. In *Proceedings of the world congress on engineering and computer science (Vol. 1)*. Newswood Limited.

YUWONO, T., FRANZ, A. and MUHIMMAH, I., 2018, April. Design of Smart Electrocardiography (ECG) Using Modified K-Nearest Neighbor (MKNN). In *2018 1st International Conference on Computer Applications & Information Security (ICCAIS) (pp. 1-5)*. IEEE.