



Implementasi Algoritma *Modified K-Nearest Neighbor* (MKNN) Untuk Klasifikasi Penerima Dana Bantuan COVID-19

Reza Chaidir¹, Hardian Oktavianto², Reni Umilasari³

Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember¹²³

Email: rezachaidir09@gmail.com^{1*}, hardian@unmuhjember.ac.id², reni.umilasari@unmuhjember.ac.id³

ABSTRAK

Bantuan Langsung Tunai (BLT) yaitu program dari pemerintahan yang pemberiannya dengan syarat dan tanpa syarat kepada masyarakat miskin dalam bentuk uang tunai atau berbagai bantuan lainnya (Kemenko Kesra, 2009). Bersamaan dengan itu, saat pandemi COVID-19 masuk ke Indonesia, BLT kembali digelar pada tahun 2020. Warga yang terkena dampak pandemi COVID-19 bukan hanya masyarakat yang terkena dampak kesehatannya, tetapi juga masyarakat yang terdampak ekonominya. Meskipun BLT dana desa di Indonesia telah dialokasikan dengan tepat, masih terdapat kekurangan, salah satunya adalah masalah sasaran penerima BLT dana desa yang belum optimal. Seperti contohnya di Desa Andongsari, Kecamatan Ambulu, Kabupaten Jember. Oleh karena itu, agar penerima BLT dana desa semakin tepat sasaran atau sesuai kriteria, maka diperlukan alat atau metode untuk membantu permasalahan tersebut, salah satu diantaranya ialah menggunakan metode klasifikasi *Modified K-Nearest Neighbor* untuk dapat mengetahui tingkat akurasi, presisi, dan recall pada penyaluran bantuan langsung tunai yang ada di Desa Andongsari. Berdasarkan hasil pengujian k yang telah dilakukan dari range 1-11 didapatkan hasil k optimal di angka 3, dan nilai akurasi sebesar 94,92%, nilai presisi sebesar 90,48%, sedangkan nilai *recall* sebesar 95%.

Kata Kunci: COVID-19, *Modified K-Nearest Neighbor*, Dana Bansos, Klasifikasi, *Data Mining*

ABSTRACT

Direct Cash Assistance (BLT), namely a government assistance program that provides conditional and unconditional assistance to the poor in the form of cash or various other assistance (Coordinating Ministry for People's Welfare, 2009). At the same time, when the COVID-19 pandemic entered Indonesia, BLT was held again in 2020. Residents affected by the COVID-19 pandemic are not only the people affected by their health, but also the people affected by their economy. Even though BLT dana desa in Indonesia has been properly allocated, there are still deficiencies, one of which is the problem of targeting beneficiaries of BLT dana desa which has not been optimal. For example, in Andongsari Village, Ambulu District, Jember Regency. Therefore, in order for Village Fund BLT recipients to be more on target or in accordance with the criteria, a tool or method is needed to help with this problem, one of which is to use the Modified K-Nearest Neighbor classification method to be able to determine the level of accuracy, precision and recall in distribution direct cash assistance in Andongsari village. Based on the results of the k test that has been carried out from the range 1-11, the optimal k results are obtained in number 3, and the accuracy value is 94.92%, the precision value is 90.48%, while the recall value is 95%.

Keywords: COVID-19, *Modified K-Nearest Neighbor*, Social Assistance Funds, Classification, *Data Mining*.

1. PENDAHULUAN

Bantuan Langsung Tunai (BLT) adalah bantuan yang diberikan oleh pemerintah kepada masyarakat miskin dalam bentuk uang tunai untuk membantu mereka menghadapi kesulitan ekonomi di tengah naiknya Bahan Bakar Minyak (BBM). Bantuan Langsung Tunai (BLT) adalah merupakan salah satu dari sekian banyak model skema dukungan sosial yang berbasis pada dukungan sosial (Atikah, 2021). Indonesia juga merupakan penyedia BLT. Mekanismenya adalah memberikan jaminan kesehatan, pendidikan, santunan tunai dan pangan. Tujuannya dibagi menjadi tiga tingkatan yaitu mendekati kemiskinan, kemiskinan, dan sangat miskin (Jamsostek Indonesia). BLT awalnya dilaksanakan mulai tahun 2005, dilanjutkan tahun 2009, dan berubah nama pada tahun 2013 jadi Bantuan Langsung Sementara Masyarakat (BLSM).

Bersamaan dengan itu, saat pandemi COVID-19 masuk ke Indonesia, BLT kembali digelar pada tahun 2020. Pandemi COVID-19 yang ada di Indonesia adalah bagian dari pandemi global *corona virus disease* (COVID-19) yang berlangsung. *Severe Acute Respiratory Syndrome Coronavirus 2* (SARS-CoV-2) adalah virus yang menyebabkan penyakit ini. Menurut Ratcliffe

(2020), kasus pertama COVID-19 di Indonesia ditemukan pada 2 Maret 2020, dengan dua individu yang diyakini berasal dari Jepang. Pada 9 April, pandemi telah menyebar ke seluruh daerah di Indonesia, yakni salah satunya provinsi Jawa Timur (Reuters, 2020).

Warga yang terkena dampak pandemi COVID-19 bukan hanya masyarakat yang terkena dampak kesehatannya, tetapi juga masyarakat yang terdampak ekonominya. Oleh karena itu sudah tepat alokasi dana desa BLT pada sebagian besar penduduk yang bekerja sebagai petani, nelayan, serta penduduk desa yang bermata pencaharian sebagai guru, pedagang atau peserta usaha kecil, menengah dan mikro, dan buruh pabrik, karena hampir semua lapisan masyarakat terdampak ekonomi dari pandemi COVID-19. Menurut Abdul Halim Iskandar, Menteri Desa Pembangunan Daerah Miskin dan Imigrasi (Mendes PDTT), semua Bantuan Langsung Tunai (BLT) dari dana desa telah diberikan kepada warga desa yang terkena dampak COVID-19. Disebutkan bahwa BLT dana desa adalah dana bantuan yang diberikan kepada warga desa yang terkena dampak pandemi COVID-19 yang berdampak pada kesehatan dan ekonomi mereka. Sejak 22 Oktober 2020, BLT dana desa telah didistribusikan kepada 7.997.269 keluarga penerima manfaat (KPM) di 74.184 desa dari total 74.957 desa di Indonesia, menurut data yang dikumpulkan oleh Kemendes PDTT.

Pembagian BLT dana desa di Indonesia memiliki kekurangan, salah satunya penerima BLT dana desa yang belum tepat sasaran. Seperti contohnya di Desa Andongsari, Kecamatan Ambulu, Kabupaten Jember. Hal ini diketahui pada saat proses observasi dan pengumpulan data penelitian melalui wawancara kepada Bapak Ikhsan yang mengatakan bahwa beberapa tetangga sekitar yang memiliki tempat tinggal tidak sesuai kriteria. Dalam artian, tempat tinggal tersebut melewati batas kriteria atau dapat dikatakan mampu namun warga tersebut masih mendapatkan BLT dana desa. Hal ini dianggap pembagian BLT dana desa belum tepat sasaran. Oleh karena itu, agar penerima BLT dana desa semakin tepat sasaran atau sesuai kriteria, maka diperlukan alat atau metode untuk membantu permasalahan tersebut. Metode digunakan yaitu dengan cara *Modified K-Nearest Neighbor* (MKNN) yang dimodifikasi. Kemampuan teknologi *data mining*, klasifikasi, memanfaatkan kesalahan masa lalu untuk meningkatkan kualitas model dan hasil analisis. (Susanto dkk., 2018). Pengkajian ini mengadopsi cara *Modified K-Nearest Neighbor* (MKNN) untuk klasifikasi pendapatan dana bantuan tingkat desa. Menerapkan algoritma MKNN untuk mengklasifikasikan penerima dana layak atau tidak berhak atas dana bantuan desa. Metode ini digunakan untuk mengklasifikasikan standar yang ditentukan pemerintah dengan data K tetangga terdekat untuk mendapatkan nama sementara penerima dana yang berhak atau tidak berhak atas subsidi desa, sehingga memudahkan pemerintah untuk mendapatkan nilai akurasi yang tinggi.

Ada modifikasi pada metode *K-Nearest Neighbor*. Metode MKNN dibagi menjadi dua proses: pembobotan *K-Nearest Neighbor*, atau *voting* bobot, dan verifikasi data pelatihan. Dibandingkan metode K-NN yang tidak melewati prosedur verifikasi data *training*, pengklasifikasian dilakukan berdasarkan bobot nilai kelas tertinggi yang diverifikasi pada k data *training* untuk mengklasifikasikan data *testing*. Metode verifikasi data *training* ini bisa memaksimalkan data *training* dengan efektifitas tinggi dan mendekati data *testing* (Parvin dkk., 2008).

2. KAJIAN PUSTAKA

A. Data Mining

Data mining merupakan suatu cara pemrosesan data yang bisa menjumpai informasi rahasia dari data. Hasil dari penggunaan metode *data mining* untuk mengolah data bisa dipergunakan untuk pengambilan keputusan di masa mendatang. *Data mining* disebut juga dengan pengenalan pola (Santoso, 2007). *Data mining* adalah metode pengolahan data berskala besar, sehingga *data mining* memegang peranan penting dalam bidang keuangan, industri, meteorologi, teknologi dan bidang lainnya. Secara umum, penelitian *data mining* menjelaskan tentang metode seperti regresi, klasifikasi, *clustering*, pemilihan variabel, dan analisis keranjang belanja (Santoso, 2007).

B. Klasifikasi

Klasifikasi merupakan tahap mencari fungsi atau model yang menjabarkan kategori data atau konsep, bertujuan agar bisa memperkirakan kategori objek dengan label yang belum diketahui. Agar tujuan ini dapat tercapai, tahap klasifikasi membentuk model yang dapat mengklasifikasikan data ke dalam kategori yang berlainan bersumber pada fungsi atau aturan tertentu. Dalam *data mining*, tahapan klasifikasi terdiri dari pembangunan model dan penerapan model (Han & Kamber, 2006).

Pada tahap pembangunan model, sebuah model dibuat untuk memecahkan persoalan klasifikasi kelas didalam data. Tahap ini adalah tahap pelatihan, menggunakan algoritma klasifikasi untuk menganalisis data pelatihan, sehingga dapat mengkomunikasikan model pembelajaran dalam bentuk aturan klasifikasi. Pada tahap penerapan model, model yang telah disusun sebelumnya dipergunakan untuk menentukan kelas data baru, yang kelasnya sebelumnya tidak diketahui. Tahap ini digunakan untuk mengetahui seberapa akurat aturan klasifikasi pada data tes. Jika model dapat diterima, aturan ini dapat diterapkan pada klasifikasi data baru.

C. Algoritma Modified K-Nearest Neighbor

Metode *K-Nearest Neighbor* yang Dimodifikasi, juga dikenal sebagai MKNN ini, adalah hasil dari modifikasi metode *K-Nearest Neighbor*. Metode MKNN terdiri dari dua tahap yaitu validasi data pelatihan adalah tahap pertama, dan pembobotan atau *voting* berat *K-Nearest Neighbor* adalah tahap kedua. Metode validasi data pelatihan ini dapat memaksimalkan data pelatihan dengan validitas tinggi dan memiliki jarak yang dekat dengan data penilaian. Dibandingkan dengan metode K-NN, yang tidak melewati proses validasi data pelatihan, klasifikasi dilakukan pada data penilaian berdasarkan bobot nilai tertinggi kelas pada k data pelatihan yang telah divalidasi dengan jarak terdekat (Parvin dkk., 2008).

Rumus perhitungan jarak yang dipergunakan untuk mendefinisikan jarak antara titik pada data *training* (p) dan titik pada data *testing* (q) ditunjukkan pada persamaan (1) (Barmer, 2007) sebagai berikut :

$$D(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (1)$$

Keterangan:

$D(p, q)$ = jarak *Euclidean*

p_i = data ke- i

q_i = data ke- i

i = 1,2,3,..., n

Jumlah titik dalam tabel yang sama untuk semua data yang ada dalam data pelatihan dihitung untuk mengetahui validitas data pelatihan. Selain itu, semua nilai fungsi data didasarkan pada apakah nilai tetangga terdekatnya pada data pelatihan sama atau tidak.

$$V(p) = \frac{1}{k} \sum_{i=1}^k S(\text{Label}(p)) (Ni(p)) \quad (2)$$

Keterangan:

k = nilai dari tetangga yang dipergunakan

$\text{Label}(p)$ = label dari data

$Ni(p)$ = label dari tetangga terdekat dengan

S = menghitung kesamaan antar satu data dengan data yan lain.

Rumus 3 adalah definisi fungsi S :

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

Keterangan:

a = kelas a pada data *training*

b = kelas lain pada data *training* selain a

Weight voting adalah tahap pada MKNN, yang merupakan hasil modifikasi dari proses K-NN. Setiap data dalam MKNN ditimbang. Untuk data pelatihan yang valid dan jarak terdekat dengan data uji, *weight voting* berguna.

$$W(i) = \text{Validitas}(i) \cdot \frac{1}{de+\alpha} \quad (4)$$

Keterangan:

$W(i)$ = Perhitungan Bobot ke i

$\text{Validitas}(i)$ = nilai validitas ke i

De = Jarak *euclidean* data *training* dan data *testing*

α = Nilai *alpha*.

3. METODE PENELITIAN

A. Studi Literatur

Pada tahap ini, eksplorasi teoritis mendalam yang relevan akan dilakukan pada masalah yang akan diteliti. Penelitian literatur berupa pemahaman metode *K-Nearest Neighbor* yang dimodifikasi yang digunakan dalam klasifikasi data rumah tangga. Jurnal ilmiah dan buku-buku tentang isu-isu terkait dipergunakan sebagai sumber literatur untuk memahami sepenuhnya tahapan klasifikasi metode modifikasi *Modified K-Nearest Neighbor*.

B. Pengumpulan Data

Data dikumpulkan dari pemerintah Desa Andongsari di Kecamatan Ambulu, Kabupaten Jember. Algoritma *Modified K-Nearest Neighbor* dipergunakan untuk mengumpulkan dan mengolah data. Dan kumpulan data tersebut memuat beberapa atribut atau karakteristik, dan karakteristik yang dipergunakan pada proses klasifikasi disesuaikan dengan variabel standar untuk menentukan status Rumah Tangga Sasaran (RTS) sebagai calon penerima dana bantuan COVID-19 yang diperoleh dari manual. Pelaksanaan penyaluran BLT dari dana desa. Kabupaten Jember pada tahun anggaran 2020. Berikut ini adalah sekumpulan variabel standar yang dipergunakan untuk menentukan status RTS, yang selanjutnya akan digunakan dalam rumah tangga sebagai proses klasifikasi atribut.

Tabel 1. Testing Atribut

No Atribut	Nama Atribut	Keterangan
Atribut 1	Luas Lantai	<8m ² /orang
Atribut 2	Jenis Lantai	Kayu murah/bambu/tanah
Atribut 3	Dinding	Tembok tanpa plester/kayu murah/rumbia/bambu
Atribut 4	Tempat buang air	Bersama orang lain/tanpa fasilitas
Atribut 5	Sumber Air minum	Dari air hujan/sungai/sumur
Atribut 6	Penerangan	Tanpa Listrik
Atribut 7	Bahan bakar	Minyak tanah/arang/kayu bakar
Atribut 8	Konsumsi	Ayam/ susu/ daging hanya 1 kali seminggu
Atribut 9	Satu stel pakaian	Satu tahun
Atribut 10	Makan	1-2 kali sehari
Atribut 11	Tidak sanggup berobat	Ke poliklinik/puskesmas
Atribut 12	Sumber penghasilan KK	Petani berlahan <500m ² , pekerjaan lain berupah <Rp. 600.000 per bulan, buruh perkebunan, buruh bangunan, buruh nelayan, buruh tani

No Atribut	Nama Atribut	Keterangan
Atribut 13	Pendidikan KK	Tidak tamat SD/tamat SD/tidak tamat SD/tidak sekolah
Atribut 14	Tidak memiliki barang mudah dijual/tabungan	Minimal Rp 500.000

Tiap data *testing*, 14 atribut di atas diklasifikasikan menjadi dua kelas yaitu rumah tangga sasaran RTS (rumah tangga sasaran) penerima dana bantuan COVID-19 dan non rumah tangga sasaran Non RTS. Untuk menjadi RTS (rumah tangga sasaran) yakni harus mencukupi 9 sampai 14 atribut untuk mendapatkan dana bantuan. Sedangkan non RTS (rumah tangga sasaran) yang tidak mendapatkan dana bantuan yakni yang kurang dari 9 dari 14 atribut.

C. Algoritma *Modified K-Nearest Neighbor*

Pada proses klasifikasi Algoritma *Modified K-Nearest Neighbor* tahap pertama yaitu dengan memasukkan data yang akan diklasifikasikan yakni data dana bantuan COVID-19 Desa Andongsari sejumlah 195 data dana penerima bantuan langsung tunai (BLT) pada tahun 2019. Selanjutnya menghitung jarak *euclidean distance*, menghitung nilai k dan menghitung validitas, kemudian memasukkan data *testing*. Selanjutnya menghitung jarak antar data *training* dengan data *testing*. Setelah itu lakukan perhitungan *weight voting*. Tahapan selanjutnya yaitu menemukan mayoritas nilai k. Langkah terakhir yakni mengeluarkan output berupa hasil akurasi, presisi dan *recall*.

4. HASIL DAN PEMBAHASAN

Data yang digunakan ialah berupa data yang bersumber dari pemerintahan Desa Andongsari Kecamatan Ambulu Kabupaten Jember, yakni berupa data penerima bantuan dana akibat COVID-19 pada tahun 2019 yang terdiri dari 195 data penduduk, dimana memiliki 14 atribut yaitu, bahan bakar, bahan lantai, bahan dinding, penerangan, sumber air minum, porsi makan per hari, tidak sanggup berobat, sumber penghasilan, pendidikan, dan tidak memiliki tabungan.

Tabel 2. Data Dana Bantuan

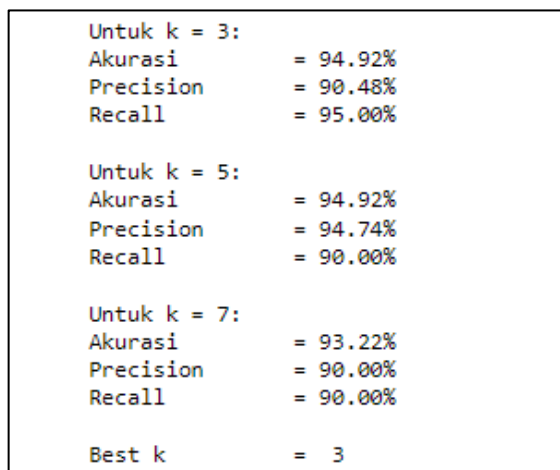
NO	Luas Lantai	Jenis Lantai	Dinding	Tempat BAB	...	Class
1	0	0	0	1	...	2
2	0	0	0	1	...	2
3	0	0	0	1	...	2
4	0	0	0	1	...	2
5	0	0	0	1	...	2
6	0	0	0	1	...	2
7	0	0	0	1	...	2
8	0	0	0	1	...	2
9	0	0	0	0	...	2
10	1	0	0	1	...	2
11	0	0	0	1	...	2
12	0	0	0	1	...	2
13	0	0	0	1	...	2
14	0	0	0	1	...	2
15	0	0	0	1	...	2
16	1	1	1	1	...	1
...
194	0	0	0	1	...	1
195	0	0	0	1	...	1

Sebelum data digunakan dalam proses pengujian, tahap persiapan dilakukan setelah pengumpulan data. Dalam proses ini akan melakukan pengindeksan kolom pada objek data

menggunakan metode *iloc* untuk memilih subset kolom dari indeks yang akan digunakan. Serta melakukan normalisasi untuk mengubah data menjadi 0 dan 1 agar lebih mudah untuk dilakukan klasifikasi.

A. Algoritma *Modified K-Nearest Neighbor*

Data yang telah disiapkan terlebih dahulu diproses memakai bahasa pemrograman *python* yang dijalankan pada *jupyter notebook* untuk diterapkan klasifikasi menggunakan metode *Modified K-Nearest Neighbor* dengan skenario *k* dari 1 sampai 11. Hasil keluaran dari perintah pada program *python* diantaranya berupa nilai *k* optimal, akurasi, presisi dan *recall*.

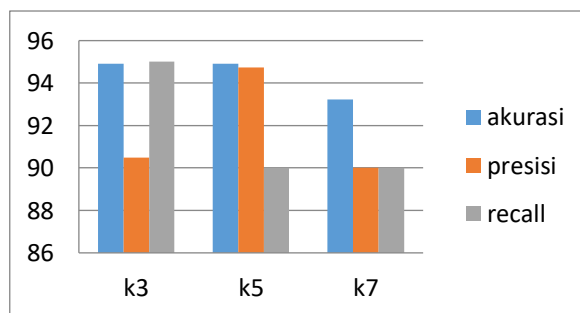


Gambar 1. Hasil Program

Implementasi metode *MKNN* menghasilkan nilai akurasi, presisi, recall dan nilai *k* optimal yang diujikan pada data penerima bantuan dana akibat COVID-19 pada tahun 2020 yang terdiri dari 195 data. Berikut nilai akurasi, presisi, recall dan nilai *k* optimal seperti pada Gambar 1.

B. Perbandingan Nilai *k*

Penting untuk melihat akurasi, presisi dan recall ini bersama-sama, terutama jika terdapat ketidakseimbangan kelas dalam dataset. Misalnya, jika *recall* rendah, berarti model cenderung melewatkan banyak data positif (*false negative*), dan mungkin perlu disesuaikan agar lebih sensitif dalam mengenali data tersebut. Selain itu, dalam beberapa kasus, seperti deteksi penyakit, *recall* mungkin lebih diutamakan daripada presisi untuk memastikan sebanyak mungkin kasus positif dapat terdeteksi, meskipun akan ada beberapa kasus positif palsu. Sedangkan di bidang lain, presisi mungkin lebih penting untuk menghindari kesalahan yang fatal akibat positif palsu.



Gambar 2. Perbandingan hasil k

5. KESIMPULAN

Berdasarkan penelitian yang dilakukan untuk mengklasifikasikan data penerima bantuan COVID-19 di Desa Andongsari pada tahun 2019 dengan menggunakan metode *Modified K-Nearest Neighbor*, maka dapat disimpulkan bahwa hasil k optimal yang diperoleh dari pengelompokan data penerima bantuan dana COVID-19 pada Desa Andongsari pada tahun 2019 menggunakan metode MKNN diperoleh nilai k optimal ialah pada $k=3$. Hasil akurasi, presisi dan *recall* yang diperoleh dari pengelompokan data penerima bantuan dana COVID-19 pada Desa Andongsari pada tahun 2019 menggunakan metode MKNN yakni nilai akurasi sebesar 94,92%, nilai presisi sebesar 90,48% dan nilai *recall* sebesar 95%.

Berdasarkan hasil penelitian yang telah dilakukan, adapun saran yang diberikan untuk penelitian selanjutnya yaitu penelitian ini dapat dijadikan bahan referensi atau dilakukan pengembangan dengan meningkatkan jumlah data atau mengganti metodenya.

6. DAFTAR PUSTAKA

- Atikah, A. K. (2021). *Implementasi Bantuan Langsung Tunai (BLT) Bagi Masyarakat Di Desa Bendosewu Kecamatan Talun Dalam Perspektif Ekonomi Syariah*. Undergraduate (S1) thesis, IAIN Kediri.
- Barmer, M. (2007). *Principles of Data Mining*. London: Springer.
- Christina, B. (2020). *Indonesia plans to relax budget deficit limit amid coronavirus*. Reuters. <https://www.reuters.com/article/us-health-coronavirus-indonesia-budget-idUKKBN21I0R4>
- Han, J., & Kamber, M. (2006). *Data Mining: Concepts and Technique, Second Edition*. Morgan Kaufmann Publishers.
- Kementerian Koordinator Bidang Kesejahteraan Rakyat. (2009). *Pelaksanaan Program Bantuan Untuk Rumah Tangga Sasaran Dalam Rangka Penanggulangan Kemiskinan*. http://www.menkokesra.go.id/sites/default/files/file_perundangan/Inpres%20Nomor%201%20Tahun%202009.pdf
- Parvin, H., Alizadeh, H. & Bidgoli, B. (2008). MKNN: Modified K-Nearest Neighbor. *Proceedings of the World Congress on Engineering and Computer Science*. San Fransisco: United Nations.
- Ratcliffe, R., & Ahmed, R. (2020). *Fears Rohingya refugees face disaster after COVID-19 reaches Cox's Bazar*. The Guardian. <https://www.theguardian.com/world/2020/may/15/fears-rohingya-refugees-face-disaster-after-COVID-19-reaches-coxs-bazar>
- Santoso, B. (2007). *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis (1 ed.)*. Yogyakarta: Graha Ilmu.
- Susanto, A., Sinaga, D., Sari, C.A., Rachmawanto, E.H., & Setiadi, D. R. I. M. (2018). A High Performace of Local Binary Pattern on Classify Javanese Character Classification. *Scientific Journal of Informatics*. Vol. 5, No. 1.