

# Classification for Covid-19 Affected Family Cash Aid Recipients Using Naïve Bayes Algorithm

Mutiara Amazona Sosiawati, Syafriandi\*, Dony Permana, Zilrahmi

Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia

\*Corresponding author: [syafriandi\\_math@fmipa.unp.ac.id](mailto:syafriandi_math@fmipa.unp.ac.id)

Submitted : 10 April 2023

Revised : 22 Mei 2023

Accepted : 31 Mei 2023

## ABSTRACT

*Village Direct Cash Assistance (BLT DD) is a kind of government financial support giving to people needed who got Covid-19 impact. There are several problems in the distribution of social assistance, one of which is recipients who are not on target or many people feel that the determination of beneficiaries is unfair. Based on the results of interviews with one of the people and one of the village officials in Tanjung Baru, Tanah Datar District, that the determination of beneficiaries was not on target, some were supposed to receive but did not receive and some did not receive but did. The beneficiaries of this assistance need to be classified and reviewed so that the people who are designated as beneficiaries are the right beneficiaries. Therefore, it is necessary to use methods to correctly classify recipients. This study uses the Naïve Bayes method to classify people who receive and do not receive aid. Based on the research that has been done, the classification using the Naïve Bayes method results in predictions that the heads of families who receive assistance and are predicted to be correct are as many as 67 people, while those who are predicted wrong are as many as 8 people, and the results of the determination of the classification of recipients of Village Fund Direct Cash Assistance in Tanjung Baru District Tanah Datar District shows a fairly high accuracy, which is equal to 89%. Meanwhile, the error rate is 11%. So, it can be concluded that the Naïve Bayes method is good enough for classification in this study. From the research results that have been obtained, it is hoped that it can be used by related parties as a reference for making policies on the distribution of aid funds.*

**Keywords:** BLT DD, Classification, Confussion Matrix, COVID-19, Naïve Bayes



This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

## I. PENDAHULUAN

Kasus pandemi COVID-19 di Indonesia memberikan dampak yang sangat besar, karena itu pemerintah telah mengambil beberapa kebijakan agar pandemi yang terjadi tidak semakin meningkat. Kebijakan yang telah diambil oleh pemerintah yaitu, melarang semua penerbangan ke dan ke China; meliburkan sekolah, kampus, serta ditutupnya sekolah, kampus, beberapa lembaga pemerintah, perusahaan swasta; dan menutup pusat hiburan (Agustino,2020). Dengan diberlakukannya kebijakan ini sangat berdampak besar pada usaha, produksi, tenaga kerja, dan pemasaran. Wabah COVID-19 ini juga telah mengakibatkan terjadinya pemutusan hubungan kerja, dan menurunkan penerimaan tenaga kerja. Semua dampak dari COVID-19 tersebut tidak hanya dirasakan oleh masyarakat perkotaan, tetapi juga dirasakan di pedesaan.

Berdasarkan intruksi dari Menteri Dalam Negeri No.3 Tahun 2020 tentang bagaimana penyelesaian masalah COVID-19 di Desa, yaitu dengan melakukan perubahan atau mempergunakan Anggaran Pendapatan Belanja Daerah (APBD) untuk Bantuan Langsung Tunai Dana Desa (BLT DD). Menurut (Hidayat,2020) dengan mempergunakan Anggaran Pendapatan Belanja Daerah untuk mengurangi dampak COVID-19 dapat diterapkan langsung di tingkat rumah tangga dan Desa.

Dalam kegiatan penyaluran bantuan terdapat beberapa permasalahan yang dapat terjadi termasuk masalah data penerima bantuan yang tidak sesuai dengan sasaran. Masalah tersebut timbul di penyaluran BLT salah satunya di kawasan Tanjung Baru Kabupaten Tanah Datar. Berdasarkan hasil wawancara dengan salah satu masyarakat dan salah satu perangkat desa di Tanjung Baru Kabupaten Tanah Datar, bahwa penetapan penerima bantuan belum tepat sasaran, ada yang harusnya menerima tetapi tidak menerima dan ada yang tidak menerima tetapi menerima. Hal tersebut menyebabkan setiap dana bantuan selesai disalurkan masih banyak masyarakat yang tidak setuju dengan daftar nama yang ditetapkan. Berdasarkan kategori status penerima bantuan tersebut yaitu menerima dan tidak

menerima, penetapan penerima bantuan tersebut perlu untuk diklasifikasikan dan ditelaah ulang supaya masyarakat yang ditetapkan sebagai penerima bantuan yang tepat.

Dengan didapatkan hasil klasifikasi dari penetapan keluarga penerima BLT DD dapat memperoleh informasi yang digunakan untuk menanggulangi masalah yang timbul pada penyaluran BLT DD, serta dapat membantu pihak yang berwenang untuk menetapkan penerima bantuan. Sehingga, tidak terjadi lagi permasalahan yang timbul pada saat penyaluran BLT DD selanjutnya.

Dalam penelitian terdahulu, metode *Naïve Bayes Classifier* digunakan untuk melihat klasifikasi dari pemilihan jurusan di SMA Negeri 1 Kampar Timur. Dengan menggunakan data training sebesar 70% dan data testing sebesar 30% didapatkan nilai akurasi yang cukup tinggi yaitu 96,19%. Sehingga untuk kasus tersebut, metode klasifikasi *Naïve Bayes* sangat cocok digunakan (Weni dkk, 2020)

Proses klasifikasi merupakan kegiatan yang bertujuan untuk mengelompokkan data pada kelas-kelas yang telah ditentukan (Han and Kamber, 2006). Berbeda dengan *cluster* yang menentukan label kelas dari tiap *cluster* yang terbentuk. Menurut Dewi (2016) terdapat beberapa metode klasifikasi, diantaranya yaitu Jaringan Saraf Tiruan, *Decision Tree*, *Naïve Bayes*, *Logistic Regression*, dan *K-NN*. Penggunaan metode klasifikasi *Naïve Bayes* akan sangat akurat bila diterapkan pada dataset yang besar (Han and Kamber, 2006). Maka dalam penelitian ini, proses klasifikasi menggunakan metode *Naïve Bayes* untuk menentukan penetapan keluarga penerima Bantuan Langsung Tunai Dana Desa di Kecamatan Tanjung Baru Kabupaten Tanah Datar.

Tujuan dari penelitian ini adalah untuk mengetahui hasil klasifikasi dari penetapan keluarga penerima Bantuan Langsung Tunai Dana Desa yang terdampak COVID-19, serta untuk mengetahui hasil akurasi dengan menggunakan metode *Naïve Bayes Classifier*. Sehingga penelitian ini dapat digunakan oleh pihak terkait sebagai acuan untuk membuat kebijakan penyaluran dana dan mengetahui sasaran masyarakat yang harus diberikan bantuan di Tanjung Baru Kabupaten Tanah Datar.

## II. METODE PENELITIAN

### A. Sumber Data Penelitian dan Variabel Penelitian

Pada penelitian ini, data yang digunakan yaitu data sekunder yang bersumber dari kantor Wali Nagari Barulak dan kantor Wali Nagari Tanjung Alam, Kecamatan Tanjung Baru Kabupaten Tanah Datar. Variabel dependen yang digunakan dalam penelitian ini yaitu Y: status penerima bantuan, dengan kategori “menerima” dan “tidak menerima”. Sedangkan variabel independen yang digunakan adalah sebanyak 7 (tujuh) variabel yaitu  $X_1$ : jenis kelamin kepala rumah tangga,  $X_2$ : pekerjaan kepala rumah tangga, dimana : 1 = petani pemilik lahan; 2 = buruh tani; 3 = pedagang makanan; 4 = lainnya,  $X_3$ : status menerima bantuan lain,  $X_4$ : kehilangan mata pencaharian/pekerjaan,  $X_5$ : memiliki anggota keluarga yang memiliki penyakit kronis/menahun,  $X_6$ : keluarga miskin yang bermukim di nagari,  $X_7$ : mempunyai anggota keluarga tunggal lanjut usia. Dengan seluruh variabel independen memiliki skala nominal. Karena seluruh variabel X pada penelitian ini hanya menunjukkan tanda atau label untuk membedakan antar variabel.

Data yang digunakan merupakan data penerima BLT DD di Kecamatan Tanjung Baru Kabupaten Tanah Datar pada Tahun 2022 dengan total data sebanyak 377 orang. Pada penelitian ini data dibagi menjadi data *training* (data latih) dan data *testing* (data uji). Hal tersebut dilakukan agar algoritma dapat lebih terlatih ketika model didapatkan dari hasil penelitian, dan jika diujikan dapat memperoleh hasil yang lebih akurat. Data *training* digunakan untuk membentuk model, sedangkan data *testing* digunakan untuk pengujian model. Pada penelitian ini menggunakan 80% data *training* dan 20% data *testing*.

### B. *Naïve Bayes Classifier*

*Naïve Bayes Classifier* adalah salah satu metode klasifikasi dalam data mining yang digunakan untuk memprediksi probabilitas keanggotaan suatu kelas. Menurut (Rish,2006) metode *Naïve Bayes* ini dapat digunakan untuk mengklasifikasikan suatu data dengan nilai akurasi yang tinggi. Teori *Bayes* pada dasarnya merupakan kemungkinan kejadian yang akan terjadi yang dapat dihitung dengan menentukan frekuensi pengalaman masa lalu, atau bisa disebut dengan teori probabilitas (Mourya,2013.11).

Berikut merupakan persamaan dari teorema *Bayes* :

$$P(C_i|X) = \frac{P(X|C_i).P(C_i)}{P(X)}$$

Dimana :

$P(C_i|X)$  : probabilitas *posterior* dimana  $C_i$  berdasarkan kondisi  $X$

$P(C_i)$  : probabilitas *prior* dari  $C_i$

$P(X|C_i)$  : probabilitas *posterior* dimana  $X$  berdasarkan kondisi  $C_i$

$P(X)$  : probabilitas *prior* dari  $X$

Untuk menghitung probabilitas yang mewakili kelas atau biasa disebut dengan probabilitas prior (sebelum masuk sampel) menggunakan persamaan :

$$P(C_i) = \frac{S_i}{s} \quad (1)$$

Dimana :

$S_i$  : jumlah data training dari kategori  $C_i$

$s$  : jumlah total data training

Tahapan klasifikasi pada metode *Naïve Bayes Classifier* menggunakan sejumlah petunjuk sebagai penentu kelas mana yang sesuai untuk sampel yang akan dianalisis, sehingga berikut merupakan penjabaran teorema Bayes :

$$P(C|X_1 \dots X_n) = \frac{P(C)P(X_1 \dots X_n|C)}{P(X_1 \dots X_n)}$$

Persamaan tersebut dapat menjelaskan peluang sampel yang telah masuk dalam kelas C (posterior) dengan karakteristik tertentu. Selanjutnya menjabarkan  $(C|X_1, \dots, X_n)$  pada teorema Bayes diatas menggunakan aturan perkalian sebagai berikut :

$$\begin{aligned} P(C|X_1, \dots, X_n) &= P(C)P(X_1, \dots, X_n|C) \\ &= P(X_1|C) P(X_2, \dots, X_n|C, X_1) \\ &= P(X_1|C)P(X_2|C, X_1)P(X_3, \dots, X_n|C, X_1X_2) \\ &= P(X_1|C)P(X_2|C, X_1), \dots, P(X_n|C, X_1, X_2, \dots, X_{n-1}) \end{aligned}$$

Dapat dilihat dari persamaan 4, bahwa hasil penambahan sampel yang telah masuk ke dalam kelas C menyebabkan kondisi yang lebih kompleks dan mempengaruhi nilai probabilitas yang tidak dapat dianalisis satu per satu.(Hayuningtyas,2017). Sehingga dapat dikatakan bahwa sangat tingginya asumsi independensi (naif), yang artinya masing-masing atribut tersebut dapat saling bebas. Oleh karena itu, berlaku persamaan berikut ini :

$$\begin{aligned} P(X_i|X_j) &= \frac{P(X_i \cap X_j)}{P(X_j)} \\ &= \frac{P(X_i)P(X_j)}{P(X_j)} \\ &= P(X_i) \end{aligned}$$

Untuk  $i \neq j$ , maka :

$$P(X_i|C, X_j) = P(X_i|C) \quad (2)$$

Dari persamaan di atas dapat diartikan bahwa asumsi independensi dapat menyederhanakan perhitungan probabilitas. Selanjutnya, penjabaran  $P(C|X_1, \dots, X_n)$  dapat disederhanakan menjadi :

$$\begin{aligned} P(C|X_1, \dots, X_n) &= P(C|X_1)P(C|X_2), \dots, P(C|X_n) \\ &= \prod_{i=1}^n P(X_i|C) \end{aligned}$$

Metode klasifikasi menggunakan *Naïve Bayes Classifier* dapat dilakukan dengan tahapan :

1. Melakukan perhitungan probabilitas prior menggunakan persamaan (1).
2. Menghitung nilai probabilitas atribut terhadap masing-masing kelas menggunakan persamaan (2)
3. Menghitung perkalian probabilitas atribut pada masing-masing kelas.
4. Menentukan nilai maksimal dari hasil perkalian probabilitas dengan probabilitas atribut.
5. Menentukan hasil klasifikasi dari data *testing*.
6. Melihat hasil ketepatan klasifikasi dengan menggunakan *confussion matrix*.

### III. HASIL DAN PEMBAHASAN

Pada metode klasifikasi *Naïve Bayes*, data yang digunakan data penerima Bantuan Langsung Tunai Dana Desa di Kecamatan Tanjung Baru Kabupaten Tanah Datar pada Tahun 2022 dengan total data sebanyak 377 orang. Dari data tersebut dibagi sebesar 80% menjadi data training dan 20% menjadi data testing.

Langkah pertama, dalam klasifikasi menggunakan *Naïve Bayes* yaitu menghitung probabilitas  $P(C_i)$  menggunakan persamaan 2. Hasil perhitungan tersebut disajikan pada Tabel 1.

**Tabel 1.** Probabilitas Status Penerima Bantuan Langsung Tunai Dana Desa (Y)

	Menerima	Tidak Menerima
	0,49	0,51

Dari Tabel 1 didapatkan hasil probabilitas status penerima Bantuan Langsung Tunai Dana Desa. Nilai peluang masyarakat menerima BLT DD adalah sebesar 0,49, sedangkan nilai peluang untuk masyarakat yang tidak menerima BLT DD adalah sebesar 0,51.

Langkah selanjutnya menghitung probabilitas  $X$  bersyarat  $C_i$  dengan menggunakan persamaan 5. Perhitungan nilai probabilitas ini menjelaskan bahwa peluang masyarakat menerima atau tidak menerima Bantuan Langsung Tunai Dana Desa berdasarkan dengan kriteria atau kondisi dari variabel  $X$ . Untuk nilai probabilitas penerima BLT DD Berdasarkan Jenis Kelamin Kepala Rumah Tangga ( $X_1$ ) disajikan pada Tabel 2.

**Tabel 2.** Probabilitas Penerima Bantuan Langsung Tunai Dana Desa Berdasarkan Jenis Kelamin Kepala Rumah Tangga ( $X_1$ )

Jenis Kelamin	Jumlah		Probabilitas $P(X C_i)$	
	Menerima	Tidak Menerima	Menerima	Tidak Menerima
Laki-laki	57	91	<b>0,385</b>	<b>0,591</b>
Perempuan	91	63	<b>0,615</b>	<b>0,409</b>
Jumlah	148	154	1	1

Dapat dilihat pada Tabel 2 nilai probabilitas menerima BLT DD yang paling tinggi yaitu kepala keluarga dengan jenis kelamin perempuan sebesar 0,615, artinya peluang untuk menerima bantuan yang paling besar yaitu kepala keluarga dengan jenis kelamin perempuan. Sedangkan untuk nilai probabilitas tidak menerima bantuan yang paling tinggi yaitu kepala keluarga dengan jenis kelamin laki-laki yaitu sebesar 0,591, artinya kepala keluarga yang berjenis kelamin laki-laki lebih besar kemungkinannya untuk tidak menerima. Maka, dapat dikatakan bahwa masyarakat yang menerima bantuan lebih banyak kepala keluarga dengan jenis kelamin perempuan dibandingkan kepala keluarga dengan jenis kelamin laki-laki.

Untuk nilai probabilitas penerima BLT DD berdasarkan pekerjaan kepala rumah tangga ( $X_2$ ) dapat dilihat pada Tabel 3.

**Tabel 3.** Probabilitas Penerima Bantuan Langsung Tunai Dana Desa Berdasarkan Pekerjaan Kepala Rumah Tangga ( $X_2$ )

Pekerjaan Kepala Rumah Tangga	Jumlah		Probabilitas $P(X C_i)$	
	Menerima	Tidak Menerima	Menerima	Tidak Menerima
Petani Pemilik Lahan	14	12	<b>0,094</b>	<b>0,078</b>
Buruh Tani	9	1	<b>0,061</b>	<b>0,006</b>
Pedagang Makanan	5	3	<b>0,034</b>	<b>0,02</b>
Lainnya	120	138	<b>0,810</b>	<b>0,896</b>
Jumlah	148	154	1	1

Dari Tabel 3 dapat dijelaskan bahwa nilai probabilitas penerima BLT DD yang tertinggi pada pekerjaan lainnya, peluang pekerjaan kepala rumah tangga yang menerima bantuan dengan syarat lainnya sebesar 0,810 dan peluang pekerjaan kepala rumah tangga yang tidak menerima bantuan dengan syarat lainnya sebesar 0,896. Artinya masyarakat penerima Bantuan Langsung Tunai Dana Desa sebagian besar berkategori pekerjaan lainnya.

Selanjutnya, untuk nilai probabilitas penerima BLT DD berdasarkan kriteria  $X_3$  (Status PKH/BPNT),  $X_4$  (Kehilangan mata pencaharian/pekerjaan),  $X_5$  (Memiliki anggota keluarga yang memiliki penyakit menahun),  $X_6$  (Keluarga miskin yang bermukim di nagari),  $X_7$  (Mempunyai keluarga tunggal lanjut usia) disajikan pada Tabel 4.

**Tabel 4.** Probabilitas Penerima Bantuan Langsung Tunai Dana Desa Berdasarkan Variabel  $X_3$  sampai dengan  $X_7$

Kriteria Menerima	$X_3$		$X_4$		$X_5$		$X_6$		$X_7$	
	Menerima	Tidak	Menerima	Tidak	Menerima	Tidak	Menerima	Tidak	Menerima	Tidak
Ya	0,034	0,454	0,54	0,032	0,567	0,039	0,223	0,136	0,263	0,006
Tidak	0,966	0,546	0,46	0,968	0,433	0,961	0,777	0,864	0,737	0,994
Jumlah	1	1	1	1	1	1	1	1	1	1

Dari Tabel 4 dapat dijelaskan untuk nilai probabilitas pada  $X_3$  (Status menerima bantuan lain) dapat diperoleh peluang masyarakat yang tidak menerima bantuan lain lebih tinggi dibandingkan yang lainnya, yaitu sebesar 0,966. Artinya, besar peluang masyarakat untuk menerima BLT DD yaitu masyarakat yang tidak menerima bantuan lain. Nilai probabilitas dengan kriteria  $X_4$  (Kehilangan mata pencaharian) yang paling tinggi yaitu 0,968. Artinya, besar peluang masyarakat tidak menerima BLT DD apabila tidak kehilangan mata pencaharian. Untuk nilai probabilitas dengan kriteria  $X_5$  (Mempunyai anggota keluarga yang memiliki penyakit menahun) yang paling tinggi yaitu 0,961. dapat dijelaskan bahwa besar peluang masyarakat tidak menerima BLT DD apabila tidak mempunyai anggota keluarga yang memiliki penyakit menahun. Untuk nilai probabilitas pada  $X_6$  (Keluarga miskin yang bermukim di nagari) yang paling tinggi yaitu 0,864. Artinya, besar peluang masyarakat tidak menerima BLT DD apabila tidak keluarga miskin yang bermukim di nagari. Sedangkan untuk nilai probabilitas pada  $X_7$  (mempunyai keluarga tunggal lanjut usia) yang paling tinggi yaitu 0,994. dapat dijelaskan bahwa besar peluang masyarakat tidak menerima BLT DD apabila tidak mempunyai keluarga tunggal lanjut usia.

Diketahui jika seseorang kepala keluarga dengan jenis kelamin perempuan, jenis pekerjaan lainnya, yang tidak menerima bantuan lain, kehilangan mata pencaharian, memiliki anggota keluarga yang memiliki penyakit menahun, tidak termasuk keluarga miskin yang bermukim di Nagari, dan tidak mempunyai anggota keluarga yang lanjut usia, Berdasarkan kondisi tersebut, akan dihitung klasifikasi kedalam kelas status penerima BLT DD menerima atau tidak menerima. Untuk menentukan nilai probabilitasnya, dapat dilihat dari Tabel 2, Tabel 3, dan Tabel 4 sebelumnya. Selanjutnya, dihitung nilai peluang dari masing-masing kategori menerima dan tidak menerima. Untuk kategori menerima didapatkan nilai peluang 0,0843. Sedangkan untuk peluang tidak menerima 0,0002. Berdasarkan peluang status penerima BLT DD menerima maupun tidak menerima, dapat dijelaskan bahwa peluang menerima lebih besar dibandingkan dengan peluang tidak menerima, Sehingga, dari salah satu data uji tersebut dapat disimpulkan bahwa status penerima BLT DD berpeluang menerima bantuan.

Berdasarkan pembagian data *training* sebanyak 302 orang dan data *testing* sebanyak 75 orang, didapatkan hasil *confussion matrix* menggunakan *software R studio* yang dapat dilihat dari Tabel 5.

**Tabel 5.** *Confussion Matrix Naïve Bayes*

$f_{ij}$	Kelas Asli (i)	Kelas Hasil Prediksi (j)	
		Kelas 1	Kelas 0
	Kelas 1	33	2
	Kelas 0	6	34

Berdasarkan Tabel 5 di atas, dapat dikatakan bahwa masyarakat yang menerima bantuan dan diprediksi menerima adalah sebanyak 33 orang, masyarakat yang tidak menerima bantuan dan di prediksi tidak menerima sebanyak 34 orang, masyarakat yang menerima bantuan dan diprediksi tidak menerima sebanyak 2 orang, dan masyarakat yang tidak menerima bantuan dan diprediksi menerima sebanyak 6 orang. Artinya, sebanyak 67 orang diklasifikasikan dengan benar dan 8 orang diklasifikasikan salah. Dari hasil yang telah diperoleh, didapatkan nilai akurasi sebesar 0,89 sedangkan laju error sebesar 0,11. Artinya, hasil akurasi yang dihasilkan dari klasifikasi status penerima BLT DD menggunakan metode *Naïve Bayes* adalah sebesar 89%, sedangkan laju error yang dihasilkan adalah 11%.

#### IV. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, klasifikasi menggunakan metode *Naïve Bayes* mendapatkan hasil prediksi sebanyak 67 orang diklasifikasikan benar, dan sebanyak 8 orang diklasifikasikan salah, dan hasil ketetapan klasifikasi penerima Bantuan Langsung Tunai Dana Desa di Kecamatan Tanjung Baru Kabupaten Tanah Datar menunjukkan akurasi yang cukup tinggi, yaitu sebesar 89%. Sedangkan untuk laju error adalah sebesar 11%. Sehingga, dapat disimpulkan bahwa metode *Naïve Bayes* cukup baik untuk klasifikasi pada penelitian ini. Dari hasil penelitian yang telah didapatkan, diharapkan bisa digunakan oleh pihak terkait sebagai acuan untuk membuat kebijakan penyaluran dana bantuan. Karena tingkat akurasi yang dihasilkan pada penelitian ini cukup tinggi, diharapkan peneliti selanjutnya menambahkan sampel untuk meningkatkan hasil akurasi dari ketetapan klasifikasi yang telah didapatkan pada penelitian sebelumnya.

#### DAFTAR PUSTAKA

- Agustino, Leo. (2020). "Analisis Kebijakan Penanganan Wabah Covid-19 : Pengalaman Indonesia". *Jurnal Borneo Administrator*, Vol.16, No.1, hal. 253-270.
- Bustami. (2014). "Penerapan Algoritma Naïve Bayes untuk Memprediksi Penjurusan Siswa di SMA Kesatrian 1 Semarang. *Jurnal Informatika*, hal 1-7.
- Chen, Hong., Hu, Songhua., Hua, Rui., & Zhao Xiuju. (2021). "Improve Naïve Bayes Classification Algorithm for Traffic Risk Management". *EURASIP Journal on Advances in Signal Processing*. Eds : Chen, et al., University of Science and Technology, Xianning, China.
- Ghozali, Imam. (2018). *Aplikasi Analisis Multivariate Dengan Program IBM SPSS 25*. Badan Penerbit Universitas Diponegoro: Semarang.
- Han, J., & M, Kamber. (2006). *Data Mining Concept and Techniques Second Edition*. San Fransisco : Morgan Kaufmann.
- Hayuningtyas, R. Y (2017). "Aplikasi Filtering of Spam Email Menggunakan Naïve Bayes". *IJCIT (Indonesian Journal on Computer and Information Technology)*, 2(1), hal 53-60.
- Hidayat, E. (2020). "Implementasi Kebijakan Dana Desa Untuk Penanggulangan Pandemi Covid-19 di Sampang". *Soetomo Communication and Humanities*, Vol. 1, No. 3.
- Hosmer, D., & Lemeshow, S. (2000). *Applied Logistic Regression. Second Edition*. USA: Jhon Wiley and Sons Inc.
- M. I. Sa'ad., D, Bryan., Kusri., Supriatin. (2020). "Decision Support System for Covid-19 Affected Family Cash Aid Recipients Using the Naïve Bayes Algorithm and the Weight Product Method. *2020 3<sup>rd</sup> International Conference on Information and Communications Technology*", hal 120-125, doi: 10.119/ICOIACT50329.2020.9331964.
- Mourya, S.K., Gupta, S. (2013). *Data Mining and Data Warehousing*. Alpha Science International Ltd.
- Rajagukguk, N., Dwi, I., & Yuciana, W.(2015). "Perbandingan Metode Klasifikasi Regresi Logistik Biner dan Naïve Bayes Pada Status Pengguna KB di Kota Tegal Tahun 2014". *Jurnal Gaussian*, Vol. 4, No. 2, hal 365-374.
- W. Xu, L. jiang. (2019). "An Attribute Value Frequency Based Instance Weighting Filter for Naïve Bayes". *Journal of Experimental and Theoretical Artificial Intelligence*, Vol. 31, No. 4, hal 225-236.
- Wanto, Anjar. (2020). *Data Mining : Alogaritma dan Implementasi*. Yayasan Kita Menulis.
- Weni S., Irwandi, Erno., & Mustakim. (2020). "Naïve Bayes Algorithm for Classification of Student Major's Specialization". *Journal of Intelligent Computing and Health Informatics*, Vol. 1, No. 1, hal 15-19.