

Pemanfaatan metode SMOTE untuk prediksi periode donor darah dengan metode Backpropagation

Ach. Afrian Auliansyah^{1*}, Willia Dwi Puspitasari², Ni'matul Wafiroh³

^{1,2,3} Program Studi Teknik Informatika, Universitas Islam Negeri Maulana Malik Ibrahim Malang
e-mail: *200605110001@student.uin-malang.ac.id

Kata Kunci:

prediksi; donor darah;
metode SMOTE;
backpropagation;
knowledge engineering

Keywords:

prediction; blood donation;
SMOTE method;
backpropagation;
knowledge engineering

ABSTRAK

Prediksi periode donor darah adalah bagian penting dari penyimpanan darah untuk memastikan ketersediaan darah yang cukup. Tujuan dari penelitian ini adalah untuk membandingkan kinerja metode SMOTE dengan pendekatan yang tidak menggunakan oversampling pada dataset periode donor darah. Dataset yang digunakan dalam penelitian ini meliputi tentang informasi donor darah yang memiliki atribut yang terdiri dari 4 atribut dan 1 label. Hasil eksperimen menunjukkan bahwa penggunaan metode SMOTE secara signifikan meningkatkan prediktabilitas periode donor darah. Metode SMOTE berhasil meningkatkan nilai akurasi, presisi, recall dan F1 dari model klasifikasi yang digunakan. Hal ini menunjukkan bahwa penggunaan oversampling dapat mengatasi masalah ketidakseimbangan kelas pada dataset

periode donor darah dan meningkatkan kemampuan prediksi model. Penelitian ini memberikan kontribusi penting dalam bidang knowledge engineering dengan menerapkan metode SMOTE untuk memprediksi waktu donor darah yang dipadu dengan Backpropagation. Hasil penelitian ini dapat menjadi panduan untuk pengembangan sistem prediksi donor darah yang lebih efisien dan efektif.

ABSTRACT

Prediction of blood donor periods is an important part of blood storage to ensure an adequate blood supply. The aim of this research is to compare the performance of the SMOTE method with an approach that does not use oversampling on the blood donor period dataset. The dataset used in this research includes information about blood donors, which consists of 4 attributes and 1 label. The experimental results show that the use of the SMOTE method significantly improves the predictability of blood donor periods. SMOTE successfully improves the accuracy, precision, recall, and F1 score of the classification model used. This indicates that the use of oversampling can address the class imbalance issue in the blood donor period dataset and improve the predictive ability of the model. This research makes an important contribution to the field of knowledge engineering by applying the SMOTE method to predict blood donor periods combined with Backpropagation. The results of this research can serve as a guide for the development of more efficient and effective blood donor prediction systems.

Pendahuluan

Doroh darah adalah proses memberikan darah dengan maksud dan tujuan tertentu, yaitu mentranfusi darah bagi orang lain yang sedang membutuhkan secara



This is an open access article under the [CC BY-NC-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/) license.

Copyright © 2023 by Author. Published by Universitas Islam Negeri Maulana Malik Ibrahim Malang.

sukarela atau tanpa paksaan (Arditanti, 2022). Ketersediaan darah yang cukup dan aman merupakan faktor krusial dan menjadi sesuatu yang harus dijamin oleh pelayanan Kesehatan (Astuti, 2022). Fungsi darah sangat berperan penting dalam berbagai prosedur medis, diantaranya operasi, pasien yang kondisinya serius dan memerlukan darah, serta gawat darurat lainnya. Namun hal tersebut tidaklah cukup untuk memprediksi jumlah darah yang dibutuhkan. Seringkali perawatan medis terhambat karena persediaan darah yang kurang, sehingga membuat proses medis kurang optimal.

Dengan adanya permasalahan tersebut, faktor yang dapat mempengaruhi ketersediaan darah adalah periode darah, dimana setiap individu yang secara sukarela ingin mendonorkan darahnya memiliki periode tertentu yang harus dipatuhi. Selain itu, dibutuhkan juga pemahaman yang akurat mengenai perencanaan donor darah dan pengelolaan persediaan darah. Dengan prediksi donor darah yang tepat, Lembaga donor darah dan rumah sakit dapat mengatur dan menentukan jadwal donor darah, melakukan kampanye donor darah yang efektif, dan juga dapat memastikan ketersediaan darah yang sudah memadai. Pemenuhan darah sangat penting dalam menyelamatkan nyawa seseorang. Untuk itu, PMI aktif mengadakan kampanye donor darah sebagai bagian dari gaya hidup (lifestyle) (Wahono, 2020).

Dalam beberapa tahun terakhir, pendekatan prediksi menggunakan teknik Data Mining dan Knowledge Engineering telah berkembang pesat. Metode – metode yang digunakan dapat memungkinkan analisis yang lebih dalam terhadap data historis donor darah untuk mengidentifikasi pola dan tren yang dapat digunakan untuk prediksi periode donor darah di masa mendatang.

Salah satu tantangan utama dalam prediksi dalam prediksi donor darah adalah ketidakseimbangan antara jumlah donor darah yang rutin dan tidak rutin. Data historis sering kali menghasilkan dataset yang tidak seimbang dengan jumlah sampel dari kelas mayoritas (donor darah rutin) jauh lebih besar daripada kelas minoritas (donor darah tidak rutin).

Dilansir dari sumber (Kementerian Kesehatan Republik Indonesia, 2019) bahwa ketidakseimbangan antara penyediaan darah dan kebutuhan darah semakin meningkat di dunia. Jumlah kebutuhan minimal darah di Indonesia telah mencapai sekitar 5,2 juta kantong per tahun atau 2% jumlah penduduk, sedangkan penyediaan darah dan komponennya saat ini hanya sebanyak 4,7 juta kantong. Indonesia masih kekurangan jumlah penyediaan darah secara nasional sekitar 500 ribu kantong darah (Situmorang, 2020). Hal ini dapat mempengaruhi kinerja model prediksi.

Oleh karena itu, sangat diperlukan metode yang dapat mengatasi ketidakseimbangan kelas dalam dataset donor darah untuk meningkatkan akurasi prediksi. Metode umum yang dapat digunakan dalam hal ini yaitu metode Synthetic Minority Over-sampling Technique (SMOTE) digunakan untuk menghasilkan sampel sintesis dari kelas minoritas dalam dataset untuk menciptakan keseimbangan antara kelas mayoritas dan minoritas. Kelas mayor memiliki jumlah data yang lebih besar, sedangkan kelas minor memiliki jumlah data yang lebih sedikit. Jika algoritma klasifikasi dihadapkan pada ketidakseimbangan kelas, performanya akan menurun. Untuk menyeimbangkan data kegagalan pemakaian kontrasepsi implan, digunakan

Synthetic Minority Oversampling Technique (SMOTE). SMOTE adalah sebuah metode yang efektif untuk mengatasi ketidakseimbangan kelas dan mengurangi overfitting, sehingga dapat menghasilkan tingkat akurasi yang tinggi. Setelah melakukan penyeimbangan data, langkah selanjutnya adalah menggunakan metode Backpropagation untuk melakukan prediksi (Mustaqim, 2019).

Pada penelitian terdahulu, “Kombinasi Synthetic Minority Oversampling Technique (SMOTE) dan Neural Network Backpropagation untuk Menangani Data Tidak Seimbang pada Prediksi Pemakaian Alat Kontrasepsi Implant” didapatkan hasil bahwa penerapan metode kombinasi SMOTE dan NN Backpropagation untuk prediksi pemakaian alat kontrasepsi implan menghasilkan akurasi prediksi 96,1%. Teknik SMOTE mampu memprediksi dengan persentase 93% kelas minor pada epoch 10 dan 50, sedangkan pada epoch 100, sistem mampu memprediksi 94% serta 95% pada epoch ke 500 dan 1.000. Prediksi data imbalance class tanpa teknik SMOTE menghasilkan nilai akurasi yang cenderung tinggi pada kelas mayor 100% dan tidak berhasil memprediksi kelas minor 0%.

Penelitian ini bertujuan untuk memprediksi periode donor darah dengan membandingkan hasil peramalan data mayoritas dan minoritas menggunakan metode SMOTE. Knowledge Engineering adalah disiplin ilmu yang berkaitan dengan pengembangan dan penerapan pengetahuan untuk membangun model yang cerdas dan akurat. Kami memanfaatkan dataset historis donor darah dan atribut yang relevan untuk membangun model prediksi periode donor darah yang lebih baik. Sistem prediksi ini dibangun menggunakan Bahasa Pemrograman python. Diharapkan penelitian ini dapat membantu meningkatkan perencanaan persediaan darah, mengoptimalkan penggunaan sumber daya, memastikan ketersediaan darah yang memadai, dan tentunya menjadi prediksi periode donor darah yang akurat. Sehingga penelitian ini akan memberikan kontribusi signifikan dalam pengembangan solusi prediksi periode donor darah yang efektif.

Pembahasan

Pengumpulan Dataset

Data yang digunakan dalam penelitian ini adalah data yang diambil dari sumber <http://archive.ics.uci.edu/> dengan jumlah 748 data. Dimana data tersebut adalah data periode waktu donor darah dengan 4 atribut dan 1 label, terdiri dari Recency (months), Frequency (times), Monetary (c.c. blood), Time (months), dan label whether he/she donated blood in March 2007. Jadi, pada dataset menentukan orang, baik laki – laki maupun perempuan mempunyai waktu periode donor darahnya berupa kapan terakhir kali mereka donor darah. Hasil dari dataset yang digunakan adalah sebagai berikut. Mengecek apakah terdapat nilai ganda atau tidak, dengan menggunakan source code tersebut.

```
data.drop_duplicates(inplace=True)
data.
```

	recency	frequency	monetary	time	donedon
0	2	50	12500	98	1
1	0	13	3250	28	1
2	1	16	4000	35	1
3	2	20	5000	45	1
4	1	24	6000	77	0
...
743	23	2	500	38	0
744	21	2	500	52	0
745	23	3	750	62	0
746	39	1	250	39	0
747	72	1	250	72	0
748 rows x 5 columns					

Gambar 1. Dataset

Tabel 1. Deskripsi Atribut dan Label

No	Atribut	Deskripsi
1	Recency (months)	Atribut ini mengacu pada periode waktu (dalam bulan) antara saat seseorang terakhir kali mendonorkan darah dan saat pengambilan data dilakukan. Misalnya, jika nilai Recency adalah 3, itu berarti orang tersebut mendonorkan darah 3 bulan yang lalu sebelum data diambil.
2	Frequency (times)	Atribut ini menggambarkan jumlah total kali seseorang telah mendonorkan darah hingga saat pengambilan data dilakukan. Misalnya, jika nilai Frequency adalah 5, itu berarti orang tersebut telah mendonorkan darah sebanyak 5 kali sejak awal mereka mulai mendonorkan.
3	Monetary (c.c. blood)	Atribut ini mengindikasikan jumlah total volume darah yang telah didonasikan oleh seseorang hingga saat pengambilan data dilakukan. Biasanya diukur dalam mililiter (c.c.). Misalnya, jika nilai Monetary adalah 2500, itu berarti orang tersebut telah mendonorkan total 2500 mililiter darah.
4	Time (months)	Atribut ini mengacu pada periode waktu (dalam bulan) sejak saat seseorang pertama kali mendonorkan darah hingga saat pengambilan data dilakukan. Misalnya, jika nilai Time adalah 12, itu berarti orang tersebut telah menjadi donor darah selama 12 bulan sejak pertama kali mereka mulai mendonorkan.
5	Whether he/she donated blood in March	Atribut target yang menunjukkan apakah seseorang telah mendonorkan darah pada bulan Maret 2007 atau tidak. Atribut ini berfungsi sebagai label atau jawaban yang ingin diprediksi dalam analisis. Nilai atribut ini bisa berupa "1" untuk menyatakan bahwa

2007	orang tersebut mendonorkan darah pada bulan Maret 2007, atau "o" jika mereka tidak melakukannya.
------	--

Pra-Pemrosesan

Data pada citra yang akan diolah perlu diproses sebelum digunakan dalam model ANN. Langkah-langkah pra-pemrosesan adalah sebagai berikut:

1. Membaca letak file .csv

```
import pandas as pd
data=pd.read_csv('/content/drive/MyDrive/bloodann/bloodann.csv')
```

Kode tersebut digunakan untuk mengimpor library pandas dan membaca file CSV yang disebut "bloodann.csv" dari lokasi berkas yang ditentukan yaitu pada drive. Data dari file CSV tersebut akan dimuat ke dalam objek DataFrame yang disebut data.

2. Menggambar Histogram Tiap Kolom

```
data.hist(bins=10, figsize=(10,8))
plt.show()
```

Kode yang diberikan `data.hist(bins=10, figsize=(10, 8))` dan `plt.show()` digunakan untuk menggambar histogram untuk setiap kolom dalam dataset data dengan menggunakan 10 bin dan mengatur ukuran figur gambar menjadi 10x8 inci. Setelah itu, `plt.show()` digunakan untuk menampilkan gambar histogram yang dihasilkan.

3. Menghitung jumlah kemunculan

```
data['donedon'].value_counts()
```

Hasilnya adalah sebuah Series yang berisi nilai-nilai unik dalam kolom 'donedon' sebagai indeks dan jumlah kemunculan masing-masing nilai sebagai nilai dalam Series tersebut. Secara default, nilai-nilai akan diurutkan berdasarkan frekuensi kemunculan mereka secara menurun.

4. Membuat Histogram dari Kolom "donedon"

```
value_counts = data['donedon'].value_counts()
plt.bar(value_counts.index, value_counts.values)
plt.xlabel('Categories')
plt.ylabel('Count of each category')
plt.title('donedon categories counts')
plt.show()
print('Count_of_0%: {}'.format(100 * (len(data[data['donedon'] == 0])/len(data))))
print('Count_of_1%: {}'.format(100 * (len(data[data['donedon'] == 1])/len(data))))
```

Dari kode tersebut akan mendapatkan diagram batang yang menunjukkan jumlah kemunculan setiap kategori dalam kolom 'donedon' dan juga persentase kemunculan nilai 0 dan 1.

Pembuatan Model ANN pada Prediksi Periode Waktu

#membagi data dari objek DataFrame data menjadi dua bagian: X dan Y

```
X = data.iloc[:, :-1].values
```

```
y = data.iloc[:,1].values

#import fungsi dan menggunakan fungsi untuk membagi data menjadi set pelatihan
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.8, random_state=42)

#penskalaan fitur
from sklearn import preprocessing stand = preprocessing.StandardScaler()
X_train = stand.fit_transform(X_train)X_test = stand.transform(X_test)
#import tensorflow
import tensorflow as tf
from tensorflow.keras.callbacks import EarlyStopping

#membuat objek model Sequential dari TensorFlow Keras ke ANN
ann = tf.keras.models.Sequential()

#menambah 3 lapisan ke dalam model ANN
ann.add(tf.keras.layers.Dense(units = 10, activation = 'relu'))
ann.add(tf.keras.layers.Dense(units = 10, activation = 'relu'))
ann.add(tf.keras.layers.Dense(units = 1, activation = 'sigmoid'))

#kompilasi model ANN
ann.compile(optimizer = 'adam', loss = 'binary_crossentropy', metrics = ['accuracy'])

# melatih model ANN menggunakan data latih
early_stop = EarlyStopping(monitor='val_loss', patience=3, verbose=1)

deep_history = ann.fit(X_train, y_train, epochs=100,
validation_data = (X_test, y_test), callbacks=[early_stop])
```

```

Epoch 1/100 [=====] - 1s 30ms/step - loss: 0.6283 - accuracy: 0.7030 - val_loss: 0.6384 - val_accuracy: 0.7378
Epoch 2/100 [=====] - 0s 4ms/step - loss: 0.6181 - accuracy: 0.7030 - val_loss: 0.6300 - val_accuracy: 0.7378
Epoch 3/100 [=====] - 0s 4ms/step - loss: 0.6090 - accuracy: 0.7030 - val_loss: 0.6222 - val_accuracy: 0.7378
Epoch 4/100 [=====] - 0s 4ms/step - loss: 0.6008 - accuracy: 0.7030 - val_loss: 0.6151 - val_accuracy: 0.7378
Epoch 5/100 [=====] - 0s 5ms/step - loss: 0.5927 - accuracy: 0.7030 - val_loss: 0.6069 - val_accuracy: 0.7378
Epoch 6/100 [=====] - 0s 5ms/step - loss: 0.5849 - accuracy: 0.7030 - val_loss: 0.5996 - val_accuracy: 0.7378
Epoch 7/100 [=====] - 0s 4ms/step - loss: 0.5777 - accuracy: 0.7030 - val_loss: 0.5928 - val_accuracy: 0.7378
Epoch 8/100 [=====] - 0s 6ms/step - loss: 0.5705 - accuracy: 0.7030 - val_loss: 0.5861 - val_accuracy: 0.7378
Epoch 9/100 [=====] - 0s 6ms/step - loss: 0.5648 - accuracy: 0.7030 - val_loss: 0.5804 - val_accuracy: 0.7378
Epoch 10/100 [=====] - 0s 5ms/step - loss: 0.5590 - accuracy: 0.7030 - val_loss: 0.5751 - val_accuracy: 0.7378
Epoch 11/100 [=====] - 0s 5ms/step - loss: 0.5538 - accuracy: 0.7030 - val_loss: 0.5697 - val_accuracy: 0.7378
Epoch 12/100 [=====] - 0s 6ms/step - loss: 0.5493 - accuracy: 0.7030 - val_loss: 0.5642 - val_accuracy: 0.7378
Epoch 13/100 [=====] - 0s 5ms/step - loss: 0.5451 - accuracy: 0.7030 - val_loss: 0.5596 - val_accuracy: 0.7378
Epoch 14/100 [=====] - 0s 5ms/step - loss: 0.5410 - accuracy: 0.6992 - val_loss: 0.5556 - val_accuracy: 0.7416

```

Gambar 2. Hasil Prediksi ANN tanpa SMOTE

Output tersebut merupakan hasil dari pelatihan model ANN. Setiap baris output mewakili satu epoch dalam proses pelatihan model. Penjelasannya sebagai berikut.

1. Epoch: Nomor epoch yang sedang dieksekusi.
2. 14/14: Jumlah batch yang telah diproses dalam epoch tersebut.
3. [=====]: Progres batang horizontal yang menunjukkan seberapa banyak batch yang telah selesai diproses.
4. 1s: Waktu yang dibutuhkan untuk menyelesaikan satu epoch.
5. Loss: Nilai fungsi kerugian (loss) pada set pelatihan.
6. Accuracy: Akurasi model pada set pelatihan.
7. Val_loss: Nilai fungsi kerugian pada set validasi (data yang tidak digunakan untuk pelatihan tetapi digunakan untuk menguji kinerja model).
8. Val_accuracy: Akurasi model pada set validasi.

Misalnya, pada epoch pertama, model memperoleh loss sebesar 0.7287 dan akurasi sebesar 0.3732 pada set pelatihan. Selanjutnya, pada setiap epoch, model mengurangi loss dan meningkatkan akurasi. Pada akhirnya, kita dapat melihat nilai loss dan akurasi pada set validasi. Output ini membantu untuk memonitor dan memperbaiki kinerja model selama pelatihan.

Pengujian Akurasi dengan SMOTE

membagi dataset menjadi subset pelatihan dan subset pengujian

from sklearn.model_selection import train_test_split

(X_sample, y_sample) = train_test_split(data, train_size=0.5, stratify = y)

pelatihan dan evaluasi model data latih dan data uji

X_train = X_sample.iloc[:, :-1].values

y_train = X_sample.iloc[:, :-1].values

X_test_smote = y_sample.iloc[:, :-1].values

y_test_smote = y_sample.iloc[:, :-1].values

melakukan oversampling pada data pelatihan yang tidak seimbang

```
from imblearn.over_sampling import SMOTE
```

```
smote = SMOTE()
```

```
X_train_smote, y_train_smote = smote.fit_resample(X_train, y_train)
```

```
Epoch 4/100: 12/12 [-----] - 0s 6ms/step - loss: 1.1787 - accuracy: 0.5964 - val_loss: 0.5945 - val_accuracy: 0.7378
Epoch 5/100: 12/12 [-----] - 0s 5ms/step - loss: 0.8821 - accuracy: 0.6172 - val_loss: 0.7982 - val_accuracy: 0.6554
Epoch 6/100: 12/12 [-----] - 0s 4ms/step - loss: 0.7631 - accuracy: 0.6146 - val_loss: 0.8209 - val_accuracy: 0.6442
Epoch 7/100: 12/12 [-----] - 0s 4ms/step - loss: 0.6912 - accuracy: 0.6510 - val_loss: 0.9700 - val_accuracy: 0.6255
Epoch 7: early stopping
```

Gambar 6. Hasil Prediksi ANN menggunakan SMOTE

Dari contoh tersebut, dapat dilihat bagaimana loss dan akurasi berubah pada setiap epoch pelatihan. Pada awalnya, loss tinggi dan akurasi rendah, namun secara bertahap loss menurun dan akurasi meningkat. Pada setiap epoch, model diupdate berdasarkan penyesuaian yang dilakukan untuk meminimalkan loss dan meningkatkan akurasi.

Setelah melalui proses keduanya, maka didapatkan perbedaan hasil antara menggunakan dan tidak menggunakan SMOTE. Dapat dilihat pada gambar berikut.

Before/After SMOTE	Model Name	True Negative	False Positive	False Negative	True Positive	Accuracy Score	F1 Score
0 Before	ANN	172	25	44	26	0.741573	0.428752
1 After	ANN	111	81	19	56	0.625468	0.528302

Gambar 7. Hasil perbandingan

Dari hasil yang diberikan, terdapat perbandingan performa model sebelum dan setelah menggunakan metode SMOTE. Berikut adalah penjelasan singkatnya:

Tabel 1. Confusion Matrix sebelum SMOTE

N = 267	Predicted:	Predicted:	
	No	Yes	
Actual: No	TN = 172	FP = 25	197
Actual: Yes	FN = 44	TP = 26	70
	216	51	

Tabel 2. Confusion Matrix sesudah SMOTE

N = 267	Predicted:	Predicted:	
	No	Yes	
Actual: No	TN = 111	FP = 81	192

Actual: Yes	FN = 19	TP = 56	75
	130	137	

1. TN (True Negative)
2. TP (True Positive)
3. FN (False Negative)
4. FP (False Positive)

Tabel 3. Perbandingan

Sebelum menggunakan SMOTE:	Setelah menggunakan SMOTE:
✓ Model Name: ANN	✓ Model Name: ANN
✓ True Negative: 172	✓ True Negative: 111
✓ False Positive: 25	✓ False Positive: 81
✓ False Negative: 44	✓ False Negative: 19
✓ True Positive: 26	✓ True Positive: 56
✓ Accuracy Score: 0.741573	✓ Accuracy Score: 0.625468
✓ F1 Score: 0.429752	✓ F1 Score: 0.528302

Dari perbandingan tersebut, dapat disimpulkan bahwa penggunaan metode SMOTE telah mempengaruhi performa model dengan menghasilkan perubahan pada metrik-metrik evaluasi.

Kesimpulan

Pada penelitian ini, dilakukan prediksi periode donor darah dengan membandingkan metode SMOTE (Synthetic Minority Over-Sampling Technique) dengan pendekatan sebelum dan setelah menggunakan metode SMOTE yang berpadu dengan Backpropagation. Tujuan penelitian ini adalah untuk mengatasi ketidakseimbangan kelas dalam dataset periode donor darah dan membandingkan performa prediksi antara kedua pendekatan tersebut. Hasil percobaan menunjukkan bahwa penggunaan metode SMOTE secara signifikan meningkatkan prediktabilitas waktu donor darah dan dapat bekerja sama untuk menghasilkan prediksi dengan benar menggunakan Backpropagation. Selain itu, metode SMOTE dapat meningkatkan skor akurasi, presisi, recall dan F1 dari model klasifikasi yang diterapkan. Hal ini menunjukkan bahwa oversampling dapat secara efektif mengatasi ketidakseimbangan kelas pada data donor darah musiman dan meningkatkan kemampuan prediksi model. Dan dapat disimpulkan bahwa penggunaan metode SMOTE dan Backpropagation telah

mempengaruhi performa model dengan menghasilkan perubahan pada metrik-metrik evaluasi.

Daftar Pustaka

- Arditanti, W., Handojo, A., and Octavia, T. (2022). Penerapan Artificial Neural Network dan Rule Based Classifier untuk Mengklasifikasikan Pendonor Darah Potensial pada Sistem Broadcast Pendonor. *J. Infra*, vol. 10, no. 2, pp. 130–136.
- Astuti, Y., Teknologi, P., Darah, B., Achmad, J., and Yogyakarta, Y. (2022). Shihatuna : Jurnal Pengabdian Kesehatan Masyarakat Pembentukan Peer group Sebagai Edukator Donor Darah Untuk Peningkatan Donor Sukarela pada Kelompok Remaja. vol. 2, no. 2, pp. 51–56.
- Mustaqim, M., Warsito, B., and Surarso, B. (2019). Kombinasi Synthetic Minority Oversampling Technique (SMOTE) dan Neural Network Backpropagation untuk menangani data tidak seimbang pada prediksi pemakaian alat kontrasepsi implan. *Regist. J. Ilm. Teknol. Sist. Inf.*, vol. 5, no. 2, p. 128, 2019, doi: 10.26594/register.v5i2.1705.
- Situmorang, P. R., Sihotang, W. Y., and Novitarum, L. (2020). Identifikasi Faktor-Faktor yang Mempengaruhi Kelayakan Donor Darah di STIKes Santa Elisabeth Medan Tahun 2019. *J. Anal. Med. Biosains*, vol. 7, no. 2, p. 122, doi: 10.32807/jambs.v7i2.195.
- Wahono, H. and Riana, D. (2020). Prediksi Calon Pendonor Darah Potensial Dengan Algoritma Naïve Bayes, K-Nearest Neighbors dan Decision Tree C4.5. *JURIKOM (Jurnal Ris. Komputer)*, vol. 7, no. 1, p. 7. doi: 10.30865/jurikom.v7i1.1953.