

PENERAPAN DATA MINING UNTUK PENENTUAN KELAYAKAN CALON PENDONOR DARAH MENGGUNAKAN METODE NAÏVE BAYES (STUDI KASUS PMI KABUPATEN KOTAWARINGIN TIMUR)

Hazanisya Priliani Hutrim¹, Nurahman²

Abstrak— Donor darah merupakan hal yang cukup penting untuk menolong seseorang yang membutuhkan darah. Dalam proses pendonoran darah, calon pendonor harus dalam kondisi yang benar-benar masuk dalam kondisi paling aman dan layak. Jika salah menentukan kriteria indikator calon pendonor darah maka akan berakibat fatal bahkan dapat menimbulkan kematian. Karena pentingnya uji kelayakan tersebut, maka dilakukan uji kelayakan pendonor menggunakan klasifikasi *data mining* dengan metode *Naïve Bayes Classifier*. Metode *Naïve Bayes Classifier* merupakan metode sederhana yang dapat digunakan untuk menentukan calon pendonor darah dengan peluang-peluang kemungkinan apakah termasuk sebagai calon pendonor atau non pendonor, yang diambil dari hasil data yang diperoleh dengan hasil perhitungan menurut rumus dasar *teorema Naïve Bayes Classification*. Kategori yang terdapat pada metode *Naïve Bayes* adalah merupakan data valid, sehingga dalam penggerjaannya sudah didapatkan hasil yang nilainya juga sudah pasti. Contohnya, pendonor darah yang ingin menyumbangkan darahnya harus memenuhi kriteria-kriteria tertentu seperti usia, Hb, tekanan darah dan berat badan. Penentuan calon donor darah diimplementasikan untuk memudahkan PMI dalam melakukan penentuan kelayakan kepada calon pendonor darah. Hasil dari penelitian ini menunjukkan bahwa akurasi model algoritma *Naïve Bayes* tergolong baik yaitu sebesar 77%.

Kata Kunci: *Data Mining, Naïve Bayes, Donor Darah, PMI.*

Abstract— Blood donation is important enough to help someone who needs blood. In the process of donating blood, prospective donors must be in conditions that are truly in the safest and most appropriate conditions. If one determines the indicator criteria for potential blood donors, it will be fatal and can even lead to death. Because of the importance of the feasibility test, a donor's feasibility test was conducted using the classification of data mining with the *Naïve Bayes Classifier* method. The *Naïve Bayes Classifier* method is a simple method that can be used to determine prospective blood donors with possible opportunities whether included as prospective donors or non donors, which are taken from the data obtained by calculating the results according to the basic formula of the *Naïve Bayes Classification* formula. The categories contained in the *Naïve Bayes* method are valid data, so the process has obtained results whose values are also certain. For example, blood donors who want to donate blood must meet certain criteria such as age, blood pressure, blood pressure and body weight. Determination of prospective blood donors is implemented to facilitate PMI in determining

eligibility for prospective blood donors. The results of this study indicate that the accuracy of the *Naïve Bayes* algorithm model is good, namely 77%.

Keywords: *Data Mining, Naïve Bayes, Blood Donors, PMI.*

I. PENDAHULUAN

1.1 Latar Belakang

Kebutuhan informasi sangatlah diperlukan seiring dengan kemajuan teknologi yang begitu pesat di zaman modern seperti pada saat ini, terutama penggunaan sistem komputerisasi yang merambat di segala bidang. Sehingga, setiap Sumber Daya Manusia (SDM) harus senantiasa dituntut untuk memenuhi keahlian, keterampilan maupun kemampuan dalam menjalankan suatu perangkat keras maupun perangkat lunak Komputer. Dengan adanya hal tersebut telah membawa masyarakat kepada perkembangan yang baik. Banyak sekali manfaat serta kemudahan yang telah dihasilkan dengan perkembangan teknologi. Penerapan teknologi informasi dapat dilihat diberbagai macam bidang seperti pendidikan, pekerjaan, dan aktivitas-aktivitas lain yang ada dalam keseharian manusia. Selain itu manusia juga membutuhkan sebuah informasi yang disajikan secara tepat, banyak cara yang dilakukan untuk memperoleh informasi salah satunya melalui internet dan untuk penyampaian informasi serta tempat untuk meletakkan informasi secara *online* didunia internet dikenal dengan sebutan *web*. Selain sebagai sarana penyaji informasi, *web* juga biasanya digunakan oleh masyarakat sebagai media untuk melakukan berbagai aktivitas.

Teknologi informasi dan komunikasi mencakup dua aspek, yaitu teknologi informasi dan teknologi komunikasi. Teknologi informasi yaitu teknologi yang mencakup segala hal yang berkaitan dengan proses, penggunaan sebagai alat bantu, manipulasi dan pengelolaan informasi. Teknologi komunikasi yaitu teknologi yang berkaitan dengan penggunaan alat bantu untuk memperoses dan mentransfer data dari perangkat yang satu kelainnya. Dari penjelasan diatas dapat diartikan bahwa teknologi informasi adalah kebutuhan manusia didalam mengambil keputusan, sebagai alat bantu dalam aktivitas sehari-hari, mengolah dan memproses informasi dalam konteks social yang

mengutungkan diri sendiri maupun masyarakat sekitar. Adapun manfaat teknologi informasi dan komunikasi sebagai berikut : [9]

1. Kemudahan dalam memperoleh informasi untuk kepentingan pendidikan, kesehatan, bisnis dan lainnya.
2. Kemudahan dalam konsultasi dan komunikasi dengan pakar atau para ahli dibidangnya dapat dilakukan dengan mudah walaupun pakar itu berada di tempat yang jauh.
3. Perpustakaan online yaitu perpustakaan dalam bentuk digital.
4. Diskusi online adalah melalui internet.
5. Inovasi dalam pembelajaran semakin berkembang dengan adanya *e-learning* yang semakin memudahkan proses pendidikan.
6. Kemajuan Teknologi informasi komunikasi juga akan memungkinkan berkembangnya kelas *virtual* yang berbasis *teleconference* yang tidak mengharuskan peserta didik berada dalam satu ruangan.
7. Sistem administrasi pada sebuah lembaga pendidikan, kesehatan, perusahaan dan lainnya akan semakin mudah dan lancar karena adanya penerapan sistem teknologi informasi komunikasi.

Palang Merah Indonesia (PMI) merupakan sebuah organisasi perhimpunan nasional di Indonesia yang bergerak dalam bidang sosial kemanusiaan, salah satu tugas dari PMI adalah melayani permintaan darah oleh masyarakat yang membutuhkan [1]. PMI didirikan oleh Pemerintah kolonial Belanda dengan nama *Het Nederland-Indische Rode Kruis* 1873. PMI menjadi salah satu organisasi kemanusiaan yang berstatus badan hukum, diundangkan dengan Undang-Undang nomor 1 tahun 2018 tentang Kepalangmerahan guna menjalankan kegiatan Kepalangmerahan sesuai dengan Konvensi Jenewa Tahun 1949, dengan tujuan untuk mencegah dan meringankan penderitaan dan melindungi korban tawanan perang dan bencana, tanpa membedakan agama, bangsa, suku bangsa, warna kulit, jenis kelamin, golongan dan pandangan politik [12].

Pelayanan darah adalah upaya pelayanan kesehatan

Darah (UDD) yang diselenggarakan oleh organisasi sosial dengan tugas pokok dan fungsinya di bidang Kepalangmerahan atau dalam hal ini Palang Merah Indonesia (PMI). Lebih lanjut, baik dalam UU No. 36/2009 tentang Kesehatan maupun Peraturan Pemerintah No.7/2011 tentang Pelayanan Darah, dinyatakan bahwa Pemerintah bertanggung jawab atas pelaksanaan pelayanan darah yang aman, mudah diakses, dan sesuai dengan kebutuhan masyarakat. Tanggung jawab pemerintah dan pemerintah daerah (Pemda) meliputi pengaturan, pembinaan, pengawasan dan pendanaan pelayanan darah untuk kepentingan pelayanan kesehatan. Sesuai penjelasan UU No. 36/2009 tentang Kesehatan Pasal 90 dan PP No. 7/2011 tentang Pelayanan Darah Pasal 46, jaminan pendanaan pemerintah diwujudkan dalam bentuk pemberian subsidi kepada UDD dari APBN, APBD dan bantuan lainnya [2].

Penyakit berasal dari bakteri, virus, maupun kuman jahat yang masuk kedalam tubuh manusia yang disebabkan antibodi manusia menurun akibat hal tertentu seperti kurangnya istirahat, kurangnya asupan gizi, lingkungan yang tidak bersih, maupun pergaulan yang tidak sehat. Penyakit memiliki cara penularan melalui beberapa cara dari saluran pernapasan, pencernaan maupun kontak fisik. Dalam penyakit mempunyai banyak nama-nama penyakit dalam bidang kesehatan begitu juga dengan obatnya dari penyakit yang diderita dapat disebarkan ataupun ditangani dengan melalui sebuah alternatif untuk mengurangi penyakit tersebut ataupun mengobatinya dapat dilakukan suatu tindakan yang diambil dengan menggunakan obat-obatan ataupun perlunya terapi untuk menangani beberapa penyakit. Oleh sebab itu perlunya sumber daya manusia dibidang kesehatan yang mampu meningkatkan keahliannya dari segi pengetahuan. Tidak hanya sumber daya manusianya saja yang dituntut tetapi pihak penyedia jasa kesehatan dari klinik, pukesmas maupun rumah sakit dituntut untuk menyediakan, melengkapi fasilitas medis yang diperlukan agar tidak ada lagi sebuah keluhan dari pasien maupun masyarakat yang disebabkan fasilitas tidak memungkinkan dan mengharuskan pasein berpindah rumah sakit dengan alasan kurangnya kelengkapan tenaga medis dan fasilitas yang ada.

Salah satu obat maupun alternatif suatu penyakit adalah darah dimana darah merupakan cairan yang terdapat pada setiap makhluk hidup (kecuali tumbuhan) darah berfungsi mengirimkan zat-zat dan oksigen yang dibutuhkan seluruh jaringan tubuh manusia dan juga berfungsi sebagai pertahanan tubuh manusia terhadap virus maupun bakteri. Darah juga mengandung berbagai banyak bahan penyusun sistem imun yang bertujuan mempertahankan sistem imun pada tubuh manusia dari berbagai penyakit. Begitu juga hormon-hormon dan dari sistem endrokin juga diedarkan melalui darah. Darah juga dapat disalurkan pada orang yang memerlukan dengan cara donor darah dan akan ditransfusikan kepada orang yang membutuhkan alternatif transfusi darah. Donor merupakan bentuk dari pemberian sesuatu pada dunia kesehatan seperti donor mata, ginjal, darah maupun organ tubuh lainnya. Donor darah salah

¹ Mahasiswa, Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Darwan Ali, Jl. Batu Berlian Kab. Kotawaringin Timur 74322 INDONESIA (telp: 0878 8447 5585; fax: (0531) 2065891; e-mail: priliani.hazanisya@gmail.com)

² Dosen, Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Darwan Ali, Jl. Batu Berlian Kab. Kotawaringin Timur 74322 INDONESIA (telp: 0878 8447 5585; fax: (0531) 2065891; e-mail: nurrahman.ikhtiar@gmail.com)

yang memanfaatkan darah manusia sebagai bahan dasar dengan tujuan kemanusiaan dan tidak untuk tujuan komersial. Peraturan Pemerintah No.7 tahun 2011 tentang Pelayanan Darah menyebutkan penyelenggaraan donor darah dan pengolahan darah dilakukan oleh Unit Donor

satunya merupakan tindakan proses atau kegiatan pemberian darah secara sukarela maupun dengan membawanya calon pendonor dimana untuk tujuan transfusi darah bagi orang yang membutuhkan. Donor darah juga termasuk salah satu alternatif terapi atau bahakan obat untuk suatu penyakit yang dimana memerlukan darah untuk transfusi pada saat situasi tertentu yang diperlukan pada saatnya maupun penyakit yang memerlukan transfusi darah secara berkala dari pendonor darah yang tetap ataupun pendonor darah sukarela dimana donor darah dapat dilakukan pada PMI terdekat pada daerah masing-masing adapun yang disebut cari bola yang dimana para petugas PMI turun langsung kelapangan untuk mencapai pendonor darah, ini dilakukan karena stok pada PMI menipis tapi permintaan darah semakin meningkat maka dari itu petugas berinisiatif untuk melakukan sosialisasi kepada masyarakat setempat guna memenuhi stok serta permintaan darah yang tersedia [10].

Pada donor darah ada beberapa penyakit yang terdapat sangat terkait pada pendonor darah contohnya anemia, leukemia dan thalasemia. Terkait donor darah thalasemia yang begitu dekat dengan pendonor darah dikarenakan penyakit ini memerlukan darah setiap waktu yang diperlukan pada pengidap-pengidapnya. Thalasemia merupakan salah satu penyakit kelainan darah yang disebabkan oleh keturunan atau genetik yang dapat berdampak pada berbagai organ akibat penyakitnya sendiri ataupun pengobatan yang diberikan pada pengidapnya. Penyakit thalasemia ini ditemukan diseluruh dunia dengan angka prevalensi gen thalasemia tertinggi terdapat pada beberapa negara tropis. Kurang lebih 3% dari penduduk dunia memiliki gen thalasemia dimana angka kejadian tertingginya sampai dengan 40% kasus adalah negara Asia. Ada juga diwilayah bagian Asia Tenggara orang-orang dengan memiliki pembawa sifat thalasemia mencapai 55juta orang. Indonesia juga merupakan salah satu negara yang memiliki penduduk pembawa thalasemia, dimana frekuensi pembawa thalasemia di Indonesia adalah sekitar 3-8%. Di beberapa tempat daerah mencapai 10%, artinya bahwa 3-8 dari 100 penduduk Indonesia merupakan pembawa gen thalassemia, dimana terdapat angka kelahiran rata-rata 23% dengan jumlah populasi penduduk yang jutaan, diperkirakan akan lahir 3000 bayi pembawa gen thalasemia di setiap tahunnya. Thalasemia ini juga merupakan salah satu penyakit yang kronis secara nyata dapat mempengaruhi kualitas hidup penderita akibat penyakit itu sendiri maupun efek terapi yang diberikan pada penderita, tidak hanya secara fisik yang dapat dilihat namun tumbuh kembang penderita juga bisa dilihat dan dikenali jika memang orang itu penderita thalasemia. Alternatif penyakit genetik thalasemia ini berupa transfusi darah terus menerus akan tetapi ada juga yang hanya melakukan transfusi darah sesekali karena berbeda jenis thalasemianya, maka dari itu perlunya stok darah yang berkecukupan agar penderita thalasemia tidak merasa kesulitan dalam mencarinya darah untuk memenuhi kebutuhan hidup mereka, sebagai yang telah diuraikan

thalasemia salah satu penyakit yang begitu membutuhkan donor darah setiap waktu [11].

Dalam rangka pemenuhan darah yang sangat penting dalam menyelamatkan nyawa seseorang terutama untuk kebutuhan penyakit thalassemia maka dari itu, PMI terus mengampanyekan donor darah sebagai bagian dari gaya hidup (*lifestyle*). Setiap tahunnya, PMI menargetkan hingga 4,5 juta kantong darah sesuai dengan kebutuhan darah nasional, disesuaikan dengan standar Lembaga Kesehatan Internasional (WHO) yaitu 2% dari jumlah penduduk untuk setiap harinya yang membutuhkan darah[3].

Pada penelitian ini dilakukan kajian terhadap PMI Kabupaten Kotawaringin Timur. Melalui data pendonor yang ada pada PMI dapat dilakukan proses *mining* untuk menentukan pola prediksi calon pendonor darah.

Dalam menentukan kelayakan calon pendonor darah, petugas Palang Merah Indonesia (PMI) melakukan *medical check up* untuk dapat mengetahui usia, kadar hemoglobin (Hb), tekanan darah dan berat badan agar memenuhi persyaratan calon pendonor darah. Dengan menggunakan metode algoritma *Naïve Bayes* untuk perhitungan data calon pendonor darah, dapat dihitung dari hasil atribut yang bernilai kontinu seperti pada atribut usia, kadar hemoglobin (Hb), tekanan darah dan berat badan, kemudian data tersebut dihitung dengan menggunakan rumus-rumus *Naïve Bayes* yang nantinya hasil dari perhitungan-perhitungan secara terperinci tersebut dapat menghasilkan suatu nilai yang lebih akurat dan efisien [4].

1.2 Rumusan Masalah

Sesuai dengan latar belakang yang telah diuraikan, maka dapat ditentukan rumusan masalah pada penelitian ini adalah bagaimana cara menentukan kelayakan pendonor darah dengan metode *Naïve Bayes* dengan menggunakan *software WEKA*?

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk menentukan kelayakan pendonor darah dengan metode *Naïve Bayes* yang akan menghasilkan layak atau tidak donor untuk calon pendonor darah dengan perhitungan yang akurat.

1.4 Manfaat Penelitian

Sesuai dengan permasalahan dan tujuan penelitian di atas, maka manfaat penelitian dapat diuraikan sebagai berikut :

- a. Mempermudah dalam pemberian informasi tentang layak atau tidaknya calon pendonor mendonorkan darahnya.
- b. Efisiensi waktu kerja dalam melakukan pemeriksaan calon pendonor.
- c. Menghasilkan informasi terkait kelayakan pendonor yang akurat dan terperinci bagi pihak yang memerlukan.
- d. Mempermudah pekerjaan pihak tenaga medis.
- e. Menambahnya referensi dalam menentukan calon pendonor darah pada PMI Kabupaten Kotawaringin Timur.
- f. Membantu meningkatkan kinerja PMI dalam menyediakan stok darah maupun memenuhi

- permintaan darah yang dibutuhkan pada rumah sakit daerah.
- g. Memaksimalkan sistem informasi dibidang kesehatan.

II. LANDASAN TEORI

2.1 Data Mining

Data mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Informasi yang dihasilkan diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau menarik dari data yang terdapat dalam basis data.

Data mining terutama digunakan untuk mencari pengetahuan yang terdapat dalam basis data yang besar sehingga sering disebut *Knowledge Discovery in Databases* (KDD). Proses pencarian pengetahuan ini menggunakan berbagai teknik-teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksikannya. Proses pencarian bersifat interaktif untuk menemukan pola atau model yang sah, baru, bermanfaat, dan dimengerti. Dalam penerapannya *data mining* memerlukan berbagai perangkat lunak analisis data untuk menemukan pola dan relasi data agar dapat digunakan untuk membuat prediksi dengan akurat [5].

Data mining mempunyai fungsi yang penting untuk membantu mendapatkan informasi yang berguna serta meningkatkan pengetahuan bagi pengguna. Tahapan proses yang umumnya dilakukan oleh *data mining* antara lain sebagai berikut : [12]

1. **Deskripsi**
Deskripsi bertujuan untuk mengidentifikasi pola yang muncul secara berulang pada suatu data dan mengubah pola tersebut menjadi aturan dan kriteria yang dapat mudah dimengerti oleh para ahli pada domain aplikasinya. Aturan yang dihasilkan harus mudah dimengerti agar dapat dengan efektif meningkatkan tingkat pengetahuan (knowledge) pada sistem. Tugas deskriptif juga merupakan tugas *data mining* yang sering dibutuhkan pada teknik *postprocessing* untuk melakukan validasi dan menjelaskan hasil dari proses data mining. *Postprocessing* merupakan proses yang digunakan untuk memastikan hanya hasil yang valid dan berguna yang dapat digunakan oleh pihak yang berkepentingan.
2. **Prediksi**
Prediksi memiliki kemiripan dengan klasifikasi, akan tetapi data diklasifikasikan berdasarkan perilaku atau nilai yang diperkirakan pada masa yang akan datang. Contoh dari tugas prediksi misalnya untuk memprediksikan adanya pengurangan jumlah pelanggan dalam waktu dekat dan prediksi harga saham dalam tiga bula yang akan datang.
3. **Estimasi**

Estimasi hampir sama dengan prediksi, kecuali variable target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun menggunakan record lengkap yang menyediakan nilai dari variable target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya di estimasi nilai dari variable target dibuat berdasarkan nilai variable prediksi. Sebagai contoh, akan dilakukan estimasi tekanan darah sistolik pada pasein rumah sakit berdasarkan umur pasein, jeniskelamin, berat badan, dan level sodium darah. Hubungan antara tekanan darah sistolik dan nilai variable prediksi dalam proses pembelajaran akan menghasilkan model estimasi.

4. **Klasifikasi**
Klasifikasi merupakan proses menemukan suatu model atau fungsi yang mendeskripsikan dan membedakan data ke dalam kelas-kelas. Klasifikasi melibatkan proses pemeriksaan karakteristik dari objek dan memasukkan objek ke dalam salah satu kelas yang sudah didefinisikan sebelumnya.
5. **Clustering**
Clustering merupakan penegelompokan data tanpa berdasarkan kelas data tertentu ke dalam kelas objek yang sama. Sebuah klaster adalah kumpulan *record* yang memiliki kemiripan suatu dengan yang lainnya dan memiliki ketidakmiripan dengan *record* dalam kluster lain. Tujuannya adalah untuk menghasilkan pengelompokan objek yang mirip satu sama lain dalam kelompok-kelompok. Semakin besar kemiripan objek dalam suatu cluster dan semakin besar perbedaan tiap cluster maka kualitas analisis klaster semakin baik.
6. **Asosiasi**
Tugas asosiasi dalam *data mining* adalah menemukannya atribut yang muncul dalam suatu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja (market basket analysis). Tugas asosiasi berusaha untuk mengungkap aturan untuk mengukur hubungan antara dua atau lebih banyak atribut.

2.2 Teknik Klasifikasi *Data Mining*

Pada referensi [13], Klasifikasi adalah penugasan data mining yang memperkerjakan sebuah objek ke salah satu kategori yang sudah didefinisikan berdasarkan atribut-atribut objeknya. Input yang dimasukan pada problem adalah sebuah dataset yang disebut sebagai *training set* yang terdiri dari jumlah sampel yang mempunyai jumlah atribut-atribut. Atribut-atribut tersebut juga bersifat berkelanjutan ketika *values* dari atribut dapat dihitung atau bersifat kategorikal ketika nilai atribut atau *values* dari atribut tidak dapat dihitung.

Salah satu atribut kategorikal disebut class-label atau atribut yang telah diklasifikasikan. Tujuannya adalah untuk menggunakan training-set dalam membangun sebuah

model dari class-label berdasarkan pada atribut-atribut lain agar model tersebut dapat digunakan untuk mengklasifikasikan data baru buka dari data-set yang sudah di-training. Teknik klasifikasi telah dipelajari secara ekstensif pada ilmu statistic, *machine-learning*, *neural networks*, dan sistem pakar [14].

2.3 Naïve Bayes

Algoritma *Naïve Bayes* merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. *Naïve Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes [12].

Teorema tersebut dikombinasikan dengan Naïve dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi Naïve Bayes diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya. Langkah-langkah algoritma *Naïve Bayes* dapat diuraikan sebagai berikut : [15]

- Baca data *training*
- Cari nilai probabilitasnya dengan cara menghitung jumlah data yang sesuai dari kategori yang sama dibagi dengan jumlah data pada kategori tersebut.
- Mendapatkan nilai dalam tabel probabilitas.

Menurut Prasetyo (2014) *Bayes* merupakan teknik prediksi berbasis probabilistik sederhana yang berdasar pada penerapan *Teorema Bayes* (atau aturan *Bayes*) dengan asumsi independensi (ketidaktergantungan) yang kuat (naïf) [6].

Prediksi *Bayes* didasarkan pada *Teorema Bayes* dengan formula umum sebagai berikut :

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)}$$

Metode ini penting karena beberapa alasan, hal ini sangat mudah untuk dibangun, tidak perlu ada yang rumit untuk parameter estimasi skema yang berulang. Ini berarti dapat segera diterapkan untuk dataset yang besar. Sangat mudah untuk menafsirkan, sehingga pengguna yang tidak terampil dalam teknologi *classifier* dapat memahami dalam membuat klasifikasi. Ini mungkin bukan *classifier* terbaik dalam setiap diberikan aplikasi, tetapi biasanya dapat diandalkan untuk menjadi kuat dan melakukan *processing data* dengan sangat baik [16].

2.4 Naïve Bayes Untuk Klasifikasi

Kaitan antara *Naïve Bayes* dan klasifikasi, korelasi hipotesis dan bukti klasifikasi adalah bahwa hipotesis dalam *Teorema Bayes* merupakan label kelas yang menjadi target pemetaan dalam klasifikasi, sedangkan bukti merupakan fitur-fitur yang menjadikan masukan dalam model klasifikasi. Jika X adalah *vector* masukan yang berisi fitur dan Y adalah label kelas, *Naïve Bayes* dituliskan dengan $P(X | Y)$ [7].

Formulasi *Naïve Bayes* untuk Klasifikasi adalah sebagai berikut :

$$P(Y|X) = \frac{P(Y) \prod_{i=1}^q P(X_i|Y)}{P(X)}$$

2.5 Deskripsi Naïve Bayes Clasification Donor Darah

- Mengelompokkan variabel berdasarkan calon pendor darah

Di dalam metode *Naïve Bayes Classification* pertama harus mengelompokkan variabel berdasarkan calon pendor darah. Pada penelitian ini terdapat beberapa variabel yaitu usia, berat badan, kadar hemoglobin, jenis kelamin, tekanan darah sistolik dan tekanan darah distolik. Variabel dalam *Naïve Bayes Classification* terdapat dua yaitu : [8]

- Variabel Diskrit (*Discrete Variable*)

Variabel Diskrit adalah variabel yang berupa data pengkategorian atau membedakan atau mengelompokkan jenis tertentu. Dari data yang diperoleh dapat diketahui bahwa terdapat 2 data diskrit dan 3 data kontinu, di antaranya :

- Jenis Kelamin
- Status Donor Darah (boleh / tidak)

- Variabel Kontinu (*Continuous Variable*)

Data kontinu adalah data yang diperoleh dari hasil penghitungan atau pengukuran, sehingga data tidak hanya berupa bilangan bulat, tetapi juga bisa dalam bentuk desimal, misalnya 3,5. Contoh dari data kontinu untuk variabel ini adalah :

- Berat Badan (kg)
- Kadar Hemoglobin (gr/dl)
- Tekanan Darah (mmHg) dan Usia (th)

- Menghitung nilai *mean* dan standar deviasi

Perhitungan *mean* dan standar deviasi hanya dapat berlaku untuk setiap *variable* yang memiliki nilai kontinu atau bisa dikatakan nilainya terus menerus saling berhubungan sesuai dengan atributnya masing-masing yaitu seperti : berat badan, usia, kadar hemoglobin, tensi atas dan tensi bawah. Perhitungan *Mean* (μ) dan Standar Deviasi (S) dapat dilihat sebagai berikut :

$$\bar{X} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Keterangan :

\bar{X} = Rata-rata hitung

x = Nilai sampel

n = Jumlah seluruh sampel

dan persamaan untuk menghitung nilai simpangan baku (standar deviasi) dapat dilihat sebagai berikut :

$$S = \sqrt{\frac{\sum(x - \bar{X})^2}{n - 1}}$$

Keterangan :

S = Standar deviasi

\bar{X} = Rata-rata hitung

x = Nilai sampel

n = Jumlah seluruh sampel

- c. Untuk nilai probabilitas dalam kategori penentuan calon pendonor darah dan probabilitas untuk setiap kategori itu sendiri, maka *teorema Bayes* dirumuskan :

$$P(E) = \frac{x}{n}$$

Keterangan :

- P = Probabilitas
 E = Event (Kejadian)
 x = Nilai sampel
 n = Jumlah seluruh sampel

- d. Perhitungan *Naïve Bayes* dari parameter-parameter calon pendonor darah akan di dapat fungsi densitas probabilitas relatif, maka fungsi densitas probabilitasnya yaitu :

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Keterangan :

- μ = mean atau nilai rata-rata (dari nilai kontinu)
 σ = Standar deviasi
 x = Nilai dari variabel pada inputan tertentu
 \exp = 2.71828

2.6 WEKA

Weka (Waikato Environment for Knowledge Analysis) merupakan aplikasi *data mining opensource* berbasis Java. Aplikasi ini dikembangkan pertama kalinya oleh suatu Universitas Waikato di Selandia Baru. Weka terdiri dari koleksi algoritma *machine learning* yang dapat digunakan untuk melakukan generalisasi / formulasi dari sekumpulan data sampling. Algoritma ini biasa juga dipanggil dari kode java kita sendiri. Weka juga memiliki tools untuk data *pre-processing, classification, regression, clustering, association rules* dan *visualization*. Weka mengorganisasikan kelas-kelas kedalam paket-paket dan setiap kelas di paket dapat mereferensi kelas lain di paket lainnya. Paket *classifiers* didalamnya terdapat implementasi hampir dari semua algoritma untuk klasifikasi dan prediksi. Kelas yang paling penting di paket ini adalah *classifier*, yang mendeklarasikan struktur umum dari skema klasifikasi serta prediksi. Kelas ini juga memiliki dua metode, yaitu build *classifier* dan *classify instance*, yang menginduk pada kelas-kelas ini. Semua kelas yang mengimplementasikan algoritma klasifikasi menginduk pada kelas *classifier*, termasuk J48. Dimana J48 yang menangani himpunan data dalam format ARFF, tidak mengandung kode untuk mengkonstruksi pohon keputusan. Kelas ini mereferensi kelas-kelas lain, kebanyakan di paket weka. *Classifiers* J48, yang mengerjakan semua proses kontruksi pohon. Weka adalah

software open source yang diterbitkan dibawah lisensi GNU (General Public License) [17].

III. METODOLOGI PENELITIAN

3.1 Objek Penelitian

Data yang diambil adalah berupa data pendonor darah yang pernah mendonorkan darahnya di Palang Merah Indonesia (PMI) Kabupaten Kotawaringin Timur, dengan ketentuan sebagai berikut :

1. Memenuhi persyaratan sebagai pendonor, syarat tersebut antara lain umur harus 18 tahun sampai dengan 60 tahun, berat badan minimal 45 kg, kadar hemoglobin minimal 12,5 – 17 g/dl, tekanan darah normal 100/10.

Apabila pendonor memenuhi persyaratan maka di perbolehkan mendonor darahnya, tetapi jika pendonor tidak memenuhi salah satu persyaratan maka tidak di perbolehkan untuk mendonorkan darahnya di PMI.

3.2 Metode Pengumpulan Data

Metode pengumpulan data yang dilakukan dalam penelitian adalah sebagai berikut :

1. Metode observasi
 Dalam metode ini penulisan melakukan observasi langsung kepada beberapa pendonor dan karyawan untuk memperoleh data yang diperlukan.
2. Metode literatur
 Dalam metode ini dilakukan dengan cara mengumpulkan data untuk memperoleh dan mencari teori-teori yang dibutuhkan untuk menyelesaikan permasalahan yang akan diteliti serta pengumpulan data dan informasi melalui buku-buku dan jurnal-jurnal.
3. Wawancara (*Interview*)
 Wawancara merupakan suatu teknik pengumpulan data dengan cara tanya jawab secara langsung kepada narasumber seperti pendonor dan karyawan yang bisa memberikan informasi atau data yang diperlukan dalam penelitian ini.

3.3 Metode Penelitian

Data yang digunakan dalam penelitian ini merupakan data primer, yaitu data calon pendonor darah di Kabupaten Kotawaringin Timur. Data tersebut diperoleh dari PMI Kabupaten Kotawaringin Timur. Variabel yang digunakan dalam penelitian ini terdiri dari status calon pendonor yaitu usia, berat badan, kadar hemoglobin, jenis kelamin, tekanan darah sistolik dan tekanan darah distolik.

Langkah-langkah yang dilakukan untuk menganalisis data penelitian adalah :

1. Mengumpulkan data calon pendonor yang akan digunakan dalam penelitian.
2. Menghitung probabilitas prior ($P(Y)$) dari data uji berdasarkan data observasi.
3. Menghitung probabilitas atribut terhadap masing-masing kelas ($P(X_i|Y)$) pada data uji berdasarkan data

- observasi menggunakan pendekatan fungsi densitas normal.
- Menghitung perkalian probabilitas prior dengan semua probabilitas atribut pada masing-masing kelas ($P(Y)$).
 - Mencari nilai maksimal dari $P(Y) \prod_{i=1}^q P(X_i|Y)$ pada kedua kelas. Nilai terbesar dari perhitungan tersebut merupakan hasil prediksi.
 - Menghitung probabilitas atribut terhadap masing-masing kelas ($P(X_i|Y)$) pada data uji berdasarkan data observasi menggunakan pendekatan selisih peluang kumulatif.
 - Menghitung perkalian probabilitas prior dengan probabilitas atribut pada masing-masing kelas ($P(Y)$).
 - Mencari nilai maksimal dari $P(Y|X) = \frac{P(Y) \prod_{i=1}^q P(X_i|Y)}{P(X)}$ pada kedua kelas. Nilai terbesar dari perhitungan tersebut merupakan hasil prediksi.
 - Pengukuran kinerja klasifikasi dengan menghitung nilai ketepatan model klasifikasi dari masing-masing pendekatan.
 - Membandingkan nilai ketepatan model klasifikasi dari masing-masing pendekatan.

IV. HASIL DAN PEMBAHASAN

4.1 Dataset

Pada penelitian ini *dataset* yang digunakan didapat dari PMI kabupaten Kotawaringin Timur, *dataset* yang digunakan adalah data dari rekap pendonor darah dimana data yang digunakan pada penelitian ini berjumlah 6919 data, maka dari itu ditampilkan pada tabel berikut ini 100 data, seperti pada Tabel 1 ini.

TABEL I.
TABEL DATA PENDONOR DARAH

Umur	Gol (Rh)	JK	Berat Badan	Tensi	HB
29	A+	Wanita	51	110/70	11,9
37	A+	Pria	64	120/80	11,6
26	B+	Pria	80	90/60	14.2
25	A+	Pria	68	140/80	13.7
33	A+	Pria	69	120/60	12,4
45	A+	Pria	74	130/80	14.2
46	A+	Pria	64	110/70	15.3
44	A+	Pria	88	130/70	16
25	A+	Wanita	93	130/90	15
42	O+	Pria	68	150/90	13.7
41	A+	Pria	80	120/80	14.4
27	A+	Pria	61	130/80	13.6
31	B+	Pria	75	160/105	14

40	B+	Wanita	91	120/70	15.4
27	O+	Wanita	75	130/80	15.9
44	O+	Wanita	78	120/80	16.5
42	O+	Pria	70	140/90	16,6
42	B+	Wanita	88	150/90	14.8
47	A+	Pria	55	120/73	12.3
39	O+	Pria	68	128/86	13.5
53	O+	Pria	76	120/73	11.8
53	A+	Pria	67	110/77	14.3
38	O+	Pria	71	137/78	15.1
34	A+	Pria	63	118/66	14.6
40	B+	Pria	67	125/87	15.2
51	A+	Pria	84	112/70	15.9
56	O+	Pria	60	115/80	13.1
48	B+	Pria	48	130/88	13.2
46	A+	Wanita	60	104/80	11
21	O+	Pria	73	135/74	14.7
28	B+	Pria	52	141/90	14.3
28	AB+	Pria	64	126/80	14.6
27	B+	Pria	80	130/100	0
38	A+	Wanita	63	120/80	14.3
32	O+	Pria	60	120/80	11
32	AB+	Pria	68	120/70	11.8
40	O+	Pria	50	120/80	12.3
26	A+	Pria	65	120/70	14.2
30	A+	Wanita	78	140/90	16.5
25	O+	Pria	74	120/90	13.7
23	O+	Pria	71	120/80	15.3
31	O+	Pria	60	110/70	13.3
40	AB+	Pria	61	120/70	14.8
28	B+	Pria	67	110/70	15.9
20	B+	Wanita	65	120/70	14.3
37	B+	Pria	61	110/60	12.3
30	B+	Pria	88	100/70	15.5
31	B+	Pria	61	120/80	15.2
36	B+	Pria	100	140/90	13.9
45	A+	Wanita	78	110/70	16.9
28	O+	Pria	70	110/80	14.7
48	B+	Pria	88	140/90	14.5
28	A+	Pria	67	105/70	11.6
32	B+	Wanita	50	120/80	14.4
46	A+	Pria	50	110/70	14.2
28	O+	Pria	62	130/90	14.1

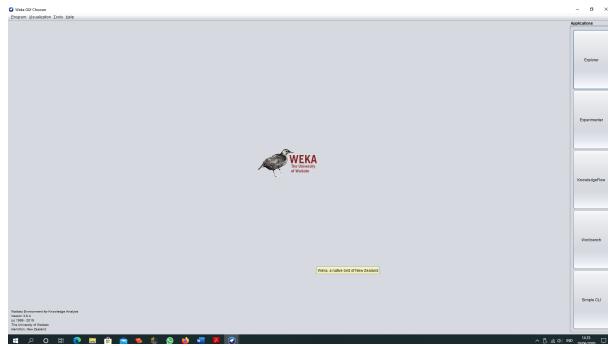
53	A+	Pria	66	120/70	12.5
38	O+	Pria	72	120/80	14.3
45	O+	Pria	79	130/90	14.1
61	B+	Pria	47	110/74	12.8
25	B+	Pria	64	112/68	12
40	A+	Wanita	75	129/89	14
42	O+	Pria	71	105/70	14.7
27	O+	Pria	66	120/60	11.3
55	B+	Pria	68	110/70	14.3
29	O+	Pria	65	100/70	11.2
23	A+	Pria	69	130/83	14
24	O+	Pria	69	108/69	14.2
38	O+	Pria	56	115/74	13.7
42	B+	Pria	64	130/66	14.7
50	O+	Pria	65	115/80	12.4
48	A+	Pria	79	142/87	14.4
45	A+	Wanita	70	128/82	16.4
45	B+	Pria	70	126/74	14.4
38	B+	Pria	73	138/86	14
40	B+	Wanita	76	136/79	14.6
41	B+	Pria	73	118/74	12.5
22	AB+	Pria	63	113/80	12.5
38	A+	Wanita	60	118/78	12.3
25	O+	Pria	65	145/96	13
21	B+	Wanita	89	114/73	15.6
30	B+	Wanita	64	127/83	13.9
19	AB+	Wanita	65	113/79	13.7
32	B+	Pria	65	129/84	12.7
25	O+	Pria	74	111/72	10.6
43	O+	Pria	68	162/102	11
43	B+	Pria	58	137/85	15.1
38	A+	Pria	55	115/59	14
50	A+	Pria	77	136/85	14.4
19	B+	Wanita	90	138/87	14.2
41	B+	Pria	65	126/90	13.6
51	O+	Pria	75	130/89	14.7
24	B+	Wanita	75	120/90	15.5
27	B+	Pria	62	143/83	14.5
32	B+	Pria	72	122/77	15.1
40	B+	Pria	55	134/81	14
42	A+	Pria	67	127/75	16.2
23	A+	Pria	64	124/67	15.1
19	O+	Wanita	65	123/79	13.3

39	A+	Pria	64	135/85	13.5
----	----	------	----	--------	------

4.2 Hasil Pengujian Menggunakan WEKA

Setelah tahap *preprocessing* dilakukan kemudian langkah selanjutnya adalah tahapan uji coba (eksperimental). Pada tahapan ini uji coba pertama yang dilakukan adalah mengimplementasi model *Naïve Bayes* dengan mengabaikan atribut yang berisikan jenis data *numerical*. Pengujian ini menggunakan *machine learning* yaitu WEKA.

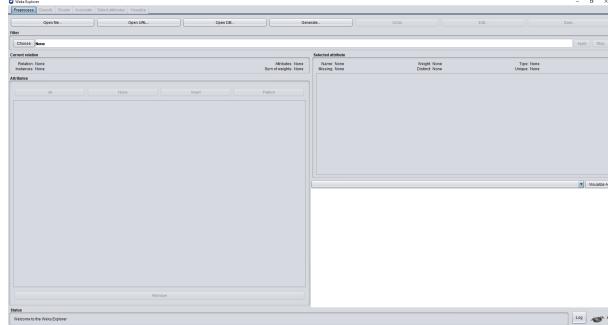
Pada gambar 1, merupakan tampilan awal pada aplikasi WEKA yang dimana terdapat menu bar Program, Visualization, Tools dan Help. Pada samping tampilan utama terdapat juga empat menu aplikasi yaitu Explorer, Experimenter, KnowledgeFlow, Workbench dan Simple CLI.



Gambar 1 Tampilan Awal WEKA

a. Langkah 1

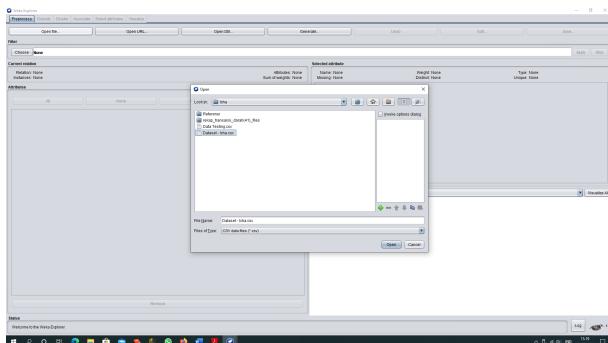
Pada tampilan awal WEKA pilih menu *Applications Explorer* sehingga akan dapat muncul tampilan window baru seperti pada gambar 2 berikut ini.



Gambar 2 Tampilan Antarmuka Aplikasi WEKA Explorer

b. Langkah 2

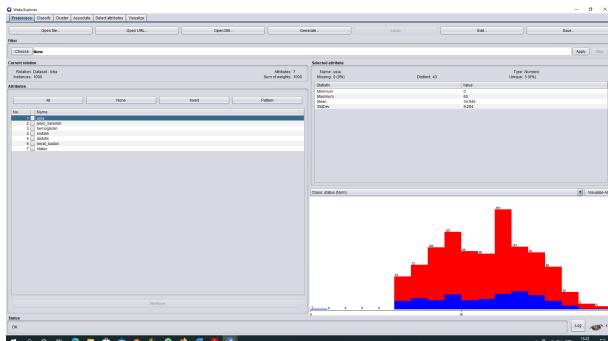
Selanjutnya klik button “Open file” dan cari tempat file *dataset* yang disimpan, yang akan digunakan sebagai data *training* dimana file *dataset* berformat csv, seperti pada gambar 3 berikut ini.



Gambar 3 Tampilan Antarmuka Open File

c. Langkah 3

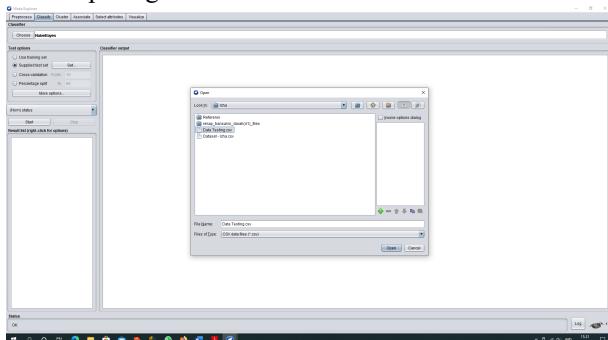
Selanjutnya akan tampil atribut-atribut yang ada di dalam tabel file dataset berformat csv. Remove atau singkirkan atribut yang tidak digunakan dalam proses *data mining* dengan cara centang pada atribut yang akan dihilangkan dan klik button “remove”, seperti pada gambar 4 berikut.



Gambar 4 Tampilan Antarmuka Pemilihan Attribute Data

d. Langkah 4

Langkah selanjutnya adalah memilih metode yang akan digunakan untuk *data mining* pendonor, contohnya di sini menggunakan metode klasifikasi, maka di tab metode pilih tab “classify” sehingga tampilan window “classify” dan di sini menggunakan klasifikasi *Naïve Bayes*, pada panel Test options pilih “Supplied Test Set” dan cari tempat file data disimpan yang akan digunakan untuk data *testing* seperti terlihat pada gambar 5 berikut ini.

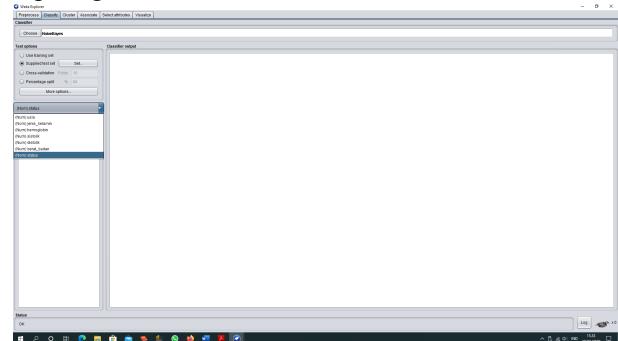


Gambar 5 Tampilan Antarmuka Pemilihan Metode dan Test Option

e. Langkah 5

Selanjutnya pilih atribut yang digunakan untuk melakukan pengklasifikasian data yang ada di bawah panel

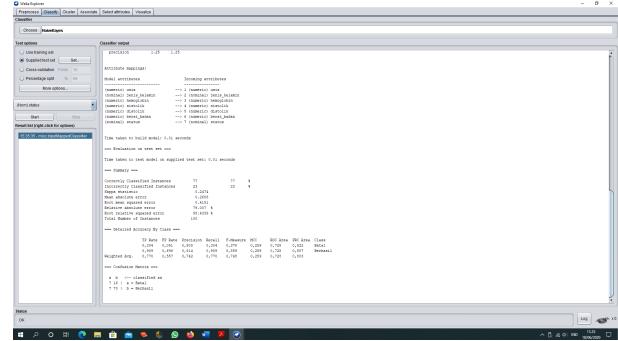
Test Options. Pada contoh ini digunakan atribut status seperti gambar 6 berikut ini.



Gambar 6 Tampilan Antarmuka Pemilihan atribut yang digunakan sebagai klasifikasi

f. Langkah 6

Selanjutnya klik button “Start” untuk menampilkan result yang dihasilkan oleh algoritma *Naïve Bayes* yang g dipilih maka akan tampil hasil dari resultnya, dapat dilihat hasilnya terdapat seperti gambar 7 berikut ini.



Gambar 7 Tampilan Antarmuka Hasil Proses Data

Tabel 2 adalah perhitungan berdasarkan data *training*, diketahui dari 100 data, 6 diklasifikasikan TIDAK LAYAK sesuai dengan prediksi yang dilakukan dengan metode algoritma *Naïve Bayes*. 6 data diprediksi TIDAK LAYAK tetapi ternyata hasilnya LAYAK, 17 data diprediksi LAYAK tetapi ternyata hasilnya TIDAK LAYAK dan 71 class LAYAK diprediksi sesuai.

TABEL II
MODEL CONFUSION MATRIX UNTUK METODE ALGORITMA NAÏVE BAYES.

Accuracy : 77,00%			
	true Tidak Layak	true Layak	class precision
pred. Tidak Layak	6	6	50.00%
pred. Layak	17	71	80.68%
class recall	26.09%	92.21%	

V. KESIMPULAN

Dalam penelitian ini dilakukan pengujian dengan model algoritma *Naïve Bayes* menggunakan *software* WEKA dari hasil evaluasi dan validasi, diketahui bahwa *Naïve Bayes* memiliki nilai akurasi cukup tinggi yaitu sebesar 77,00%. Dengan demikian, metode *Naïve Bayes* yang digunakan pada penelitian ini merupakan metode yang cukup baik dalam menentukan calon pendonor darah secara lebih efektif dan efisien.

VI. UCAPAN TERIMA KASIH

Mengucapkan terimakasih kepada seluruh pihak yang terlibat dalam penyelesaian jurnal ini, terutama kepada PMI kabupaten Kotawaringin Timur yang telah memberikan data, ijin serta tempat untuk penelitian.

DAFTAR PUSTAKA

- [1] PMI, "Sejarah PMI," Palang Merah Indonesia, 2011. [Online]. Available: [http://www.pmi.or.id/index.php/tentang-kami/sejarahpmi.html?showall=1&limitstart=". \[Accessed: 10-Nov-2019\].](http://www.pmi.or.id/index.php/tentang-kami/sejarahpmi.html?showall=1&limitstart=)
- [2] Kemenkumham RI, "Peraturan Pemerintah No. 7 Tahun 2011," 2011.
- [3] Pusdatin Kemkes RI, "Pelayanan Darah Di Indonesia," Kementerian Kesehatan RI, 2014.
- [4] R. Dewi, M. Kartasurya and A. Mawarni, "Analisis Kebijakan Donor Darah Dan Implementasi Program Rekrutmen Donor Di Unit Donor Darah (UDD PMI) Kota Pontianak," *Jurnal Manajemen Kesehatan Indonesia*, vol. 4, no. 2, pp. 109-117, 2017.
- [5] M. Ridwan, H. Suyono and M. Sarosa, "Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naïve Bayes Classifier," *Jurnal EECCIS*, vol. 7, no. 1, pp. 59-64, 2013.
- [6] A. Jananto, "Algoritma Naïve Bayes untuk Mencari Perkiraan Waktu Studi Mahasiswa," *Jurnal Dinamik*, vol. 18, no. 1, 2013.
- [7] A. Mukminin and D. Riana, "Komparasi Algoritma C4. 5, Naïve Bayes Dan Neural Network Untuk Klasifikasi Tanah," *Jurnal Informatika*, vol. 4, no. 1, 2017.
- [8] Kurniawan, A., (2010), Penentuan Calon Pendonor Darah Menggunakan Algoritma Naïve Bayes Classification, Universitas Dian Nuswantoro, Semarang.Group EUROASPIREIIS: Lifestyle and risk management and use of drug therapies in coronary patients from 15 countries.
- [9] <https://www.gurupendidikan.co.id/teknologi-informasi-dan-komunikasi/> diakses tanggal 12 Agustus 2020.
- [10] T.M.Z Castaka Agus Sugiarto, "Rancang bangun aplikasi donor darah berbasis mobile di PMI Kabupaten Bandung" *Jurnal Ilmiah Manajemen Informatika dan Komputer*, 2017.
- [11] R, Pranajaya, Nurchairina, "Faktor yang berhubungan dengan kualitas hidup anak thalasemia", *Jurnal Keperawatan*, volume XII, No. 1, April 2016.
- [12] Hermanto Wahono, Dwiza Riana "Prediksi calon pendonor darah potensial dengan algoritma naïve bayes, k-nearest neighbors dan decision tree C4.5", *Jurnal Riset Komputer*, Vol. 7 No. 1, Februari 2020.
- [13] J. Han and M. P. J. Kamber, *Data Mining:Concept and Techniques*, Burlington: Elsevier, 2012.
- [14] Mitchell, *Machine Learning*, 1 Ed, Indian: Mcgraw Hill, 2013.
- [15] Bustami, "Penerapan Algoritma Naïve Bayes," *J. Inform*, 2014.
- [16] S. Dewi, "Komparasi 5 Metode Algoritma Klasifikasi Data Mining Pada Prediksi Keberhasilan Pemasaran Produk Layanan Perbankan," *None*, Vol. 13 No. 1, pp. 60-66, 2016.
- [17] Ketut Jaya Atmaja, Ida Bagus Gede Anandita and Ni Kadek Ceryna Dewi, "Penerapan Data Mining Untuk Memprediksi Potensi Pendonor Darah Menjadi Pendonor Tetap menggunakan Metode Decision Tree C4.5," *Jurnal S@CIES Volume 7*, Nomor 2, April 2017.