



## Deteksi dini resiko tuberkulosis di kota ternate: pelacakan dan implementasi algoritma klasifikasi

Andi Sitti Nur Afiah<sup>1</sup>, Soesanti<sup>2</sup>, Abd Hakim Husen<sup>3</sup>, Firman Tempola\*<sup>4</sup>

Email: <sup>1</sup>nurafiahnasir@yahoo.com, <sup>2</sup>drsoesanti73@gmail.com, <sup>3</sup>abdhakim@unkhair.ac.id, <sup>4</sup>firman.tempola@unkhair.ac.id

<sup>1,2,3</sup>Program Studi Pendidikan Dokter, Fakultas Kedokteran, Universitas Khairun

<sup>4</sup>Program Studi Teknik Informatika, Fakultas Teknik, Universitas Khairun

Diterima: 19 Agustus 2022 | Direvisi: 23 Agustus 2022 | Disetujui: 24 Agustus 2022

©2020 Program Studi Teknik Informatika Fakultas Ilmu Komputer,  
Universitas Muhammadiyah Riau, Indonesia

### Abstrak

Tingkat Kematian akibat virus Tuberkulosis masih cukup tinggi. sebagaimana dilaporkan di Kota Ternate Provinsi Maluku Utara bawah pada tahun 2018 sebanyak 452 kasus / 100.000 penduduk. Dengan tingkat kematian mencapai 23 orang / 100.000 penduduk. Tingkat kematian yang begitu tentu harus ada Langkah-langkah preventif sehingga dapat mengurangi resiko kematian akibat dari penyakit TBC. Untuk perlu dilakukan proses pelacakan kepada pasien suspek TBC di kota ternate. Pada penelitian ini menggunakan sampel dari wilayah kerja Puskesmas Kalumata Kota Ternate. Dengan jumlah sampel yang digunakan sebanyak 100 sampel didapat 47% pasien beresiko TBC. Dimana 70% didominasi kaum laki-laki. Selanjutnya data-data yang telah dianalisis oleh dokter, selanjutnya dilakukan proses klasifikasi dengan menggunakan dua metode klasifikasi yaitu metode Support Vector Machine (SVM) dan Jaringan Saraf Tiruan. Namun sebelum diterapkan metode klasifikasi, terlebih dahulu dilakukan proses imputasi untuk penanganan missing value. Dalam penelitian digunakan imputasi modu. Hasil pengujian yang dilakukan didapat akurasi tertinggi untuk metode SVM sebesar 92,5%, sedangkan ketika menerapkan jaringan saraf tiruan didapat akurasi tertinggi sebesar 91,66%. Namun saat diterapkan proses validasi dengan menggunakan k-fold cross validasi didapatkan rata-rata akurasi tertinggi yaitu 85,08 % dengan menggunakan 3-fold dan algoritma yang diterapkan adalah jaringan saraf tiruan

**Kata kunci:** tuberkulosis, pelacakan, SVM, jaringan saraf tiruan

### *Early detection of the risk of tuberculosis in the city of ternate: tracking and implementation of classification algorithm*

#### **Abstract**

*The death rate from the Tuberculosis virus is still quite high. as reported in Ternate City, North Maluku Province, in 2018 as many as 452 cases / 100,000 population. With a death rate of 23 people / 100,000 population. Such a death rate must have preventive measures so as to reduce the risk of death due to TB disease. It is necessary to carry out a tracking process for suspected TB patients in the city of Ternate. In this study using samples from the working area of the Kalumata Health Center, Ternate City. With the number of samples used as many as 100 samples obtained 47% of patients at risk of tuberculosis. Where 70% is dominated by men. Furthermore, the data that has been analyzed by the doctor, then the classification process is carried out using two classification methods, namely the Support Vector Machine (SVM) method and the Artificial Neural Network. However, before applying the classification method, the imputation process is carried out first for handling missing values. In this study, the mode of imputation was used. The results of the tests carried out obtained the highest accuracy for the SVM method of 92.5%, while when applying artificial neural networks the highest accuracy was obtained at 91.66%. However, when the validation process is applied using k-fold cross validation, the highest average accuracy is 85.08% using 3-fold and the algorithm applied is an artificial neural network.*

**Keywords:** tuberkulosis, tracking, SVM, artificial neural network

## 1. PENDAHULUAN

Salah satu penyakit yang tingkat kematian masih cukup tinggi di dunia adalah penyakit yang diakibatkan oleh *Mycobacterium* atau yang biasanya dikenal dengan tuberkulosis. Tuberkulosis merupakan salah satu jenis bakteri basil, dimana untuk pengobatannya membutuhkan waktu yang agak lama. Di negara-negara berkembang Tuberkulosis masih menjadi masalah kesehatan hingga saat ini. Meskipun saat ini vaksin dari tuberkulosis atau yang dikenal *Bacillus Calmette Guering* (BCG) terus digalakkan namun penyakit TBC masih belum terselesaikan [1].

Laporan World Health Organization menyebutkan dalam Global Surveillance pada tahun 2019-2020 bahwa penyakit tuberkulosis di dunia mengalami penurunan. Meski demikian jumlah penderita masih diprediksi sekitar 10 juta jiwa. Dengan tingkat kematian diperkirakan mencapai 1,2 juta [2]. Infeksi yang dialami manusia yang diakibatkan oleh *Mycobacterium* yaitu pada organ paru. Sehingga sering juga disebutkan TB paru (Tuberkulosis Paru). Cara penularan bakteri ini juga terbilang cukup cepat hal ini karena bakteri ini dapat menular melalui udara yaitu lewat batuk atau bersin dari penderita.

Di Indonesia pada tahun 2017 dilaporkan Jumlah kasus baru tuberkulosis sebanyak 420.994. Dari kasus tersebut lebih banyak di alami oleh laki-laki yaitu sebanyak 1,4 kali lebih besar dibandingkan pada perempuan. Bahkan berdasarkan Survei Prevalensi Tuberkulosis prevalensi pada laki-laki 3 kali lebih tinggi dibandingkan pada perempuan. Sama halnya juga yang di alami oleh negara-negara lainnya. Hal ini diakibatkan karena pola hidup dari laki-laki yang kurang memperhatikan seperti merokok dan kurangnya ketidapatuhan minum obat. Survei ini menemukan bahwa dari seluruh partisipan laki-laki yang merokok sebanyak 68,5% dan hanya 3,7% partisipan perempuan yang merokok. Sedangkan Indonesia pada tahun 2020 dalam laporan tersebut penderita tuberkulosis mencapai 845.000 dan 67% merupakan kasus baru [3].

Tuberkulosis merupakan penyebab kematian kedua setelah Human Immunodeficiency Virus. untuk itu sangat penting kepatuhan dari pasien dalam pengobatan TB Paru agar mencapai tujuan Millenium Development Goals (MDGs) dalam rangka menurunkan penyebaran TB. Di Maluku Utara tingkat penderita penyakit TBC juga terus mengalami peningkatan sebagaimana dilaporkan pada tahun 2016 sebanyak 134 kasus sedangkan pada tahun 2017 Case Notification Rate (CNR) capai 159. Di kota Ternate dilaporkan pada tahun 2018 sebanyak 452 kasus / 100.000 penduduk. Dengan tingkat kematian mencapai 23 orang / 100.000 penduduk [4].

Tingkat kematian yang begitu tinggi, tentu perlu ada langkah-langkah nyata untuk dapat mengurangi tingkat penderita TBC. Salah satu caranya yaitu dengan proses pelacakan untuk deteksi awal resiko TBC pada pasien suspek. Deteksi awal ini berdasarkan pola hidup dari setiap pasien, sebagaimana yang akan dilakukan pada pasien yang berada pada wilayah kerja Puskesmas Kalumata Kota Ternate. Untuk mengetahui resiko dan tidak beresiko pasien diminta untuk mengisi kuesioner dari dokter setempat, kemudian dokter mengambil keputusan mana pasien yang beresiko dan mana yang tidak beresiko.

Di era 4.0 seperti ini data-data yang telah dianalisis oleh dokter kemudian bisa dimanfaatkan dengan menerapkan algoritma machine learning untuk klasifikasi seperti Support Vector Machine, Jaringan Saraf Tiruan, Naïve Bayes maupun Decision Tree. Beberapa penelitian dengan menerapkan metode machine learning untuk proses klasifikasi pada data Tuberculosis pernah dilakukan sebelumnya diantaranya [5] dengan menerapkan metode naïve bayes, dalam penelitian tersebut kinerja sistem berupa nilai akurasi sebesar 92%. [6] dengan menggunakan metode yang sama yaitu naïve bayes akurasi sistem mencapai 94,18%. [7] Menerapkan Classification Regressino Tree didapat akurasi sistem sebesar 83,7%.

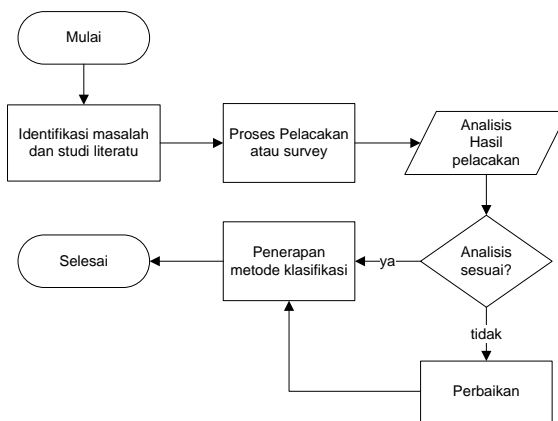
Penelitian ini akan menerapkan beberapa metode supervised learning seperti support vector machine dan jaringan saraf tiruan. Penggunaan metode supervised learning untuk penyakit TBC pernah diterapkan [8] dengan metode yang digunakan yaitu support vector machine untuk memprediksi jumlah penderita Tuberkulosis. [9] Menerapkan SVM untuk penyakit tuberculosis namun pada penelitian tersebut belum dilakukan proses validasi. Sedangkan metode jaringan saraf tiruan pernah dilakukan oleh [10] memprediksi penyakit TBC berdasarkan citra Rontgen dengan hasil akurasi yang didapat sebesar 79,41%.

Penerapan SVM juga pernah diterapkan pada kasus lain sebagaimana yang dilakukan oleh [11] pada klasifikasi pemakain minyak goreng. Begitupun dengan penerapan Jaringan saraf tiruan sebagaimana yang dilakukan [12][16] pada pengenalan sidik jari untuk smart home.

Penelitian ini selain menerapkan beberapa metode supervised learning, juga melakukan validasi dengan k-fold cross validasi. Penggunaan k-fold cross validasi pernah diterapkan [13] pada data-data aktifitas gunung berapi. kemudian akan dievaluasi dengan beberapa teknik evaluasi metode klasifikasi seperti akurasi, presisi, recall dan f1-Score sebagaimana yang pernah dilakukan [14, 15].

## 2. METODE PENELITIAN

Penelitian ini bersifat kualitatif dimana data-data yang diterapkan yaitu berupa data-data hasil pelacakan dari pasien suspek tuberculosis. Selanjutnya dilakukan proses eksperimen dengan menerapkan beberapa algoritma yang ada di machine learning yaitu metode klasifikasi. Kemudian akan dilakukan uji kinerja algoritma dengan menggunakan beberapa metode untuk menguji kinerja algoritma klasifikasi seperti uji akurasi, uji presisi dan recall maupun uji F1-Score. Adapun Langkah-langkah dalam penelitian ini yaitu ditunjukkan pada Gambar 1.



Gambar 1. Langkah-langkah penelitian

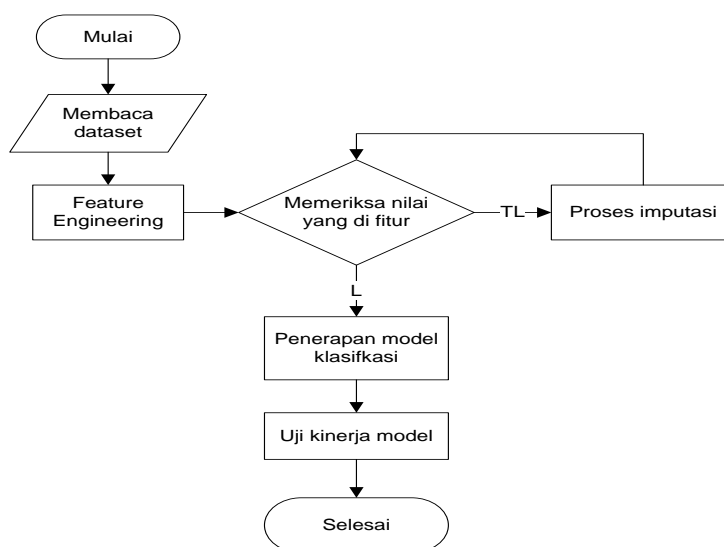
1. Penelitian ini dimulai dengan proses identifikasi permasalahan dan dilanjutkan dengan studi literatur terhadap referensi yang sesuai dengan lingkup permasalahan penelitian.
2. Proses berikutnya adalah penyusunan kuesioner dan dilanjutkan dengan proses pelacakan atau survey pada objek penelitian.
3. Hasil pelacakan kemudian dilakukan analisis statistik untuk didapatkan insight dari penelitian
4. Jika hasil analisis sudah sesuai maka diterapkan salah satu metode machine learning yaitu metode klasifikasi, namun jika belum sesuai maka akan dilakukan perbaikan analisis hingga sesuai baru diterapkan metode klasifikasi

2.1. Populasi dan Sampel

Populasi dalam penelitian ini merupakan pasien suspek TBC di wilayah kerja Puskesmas Kalumata Kota Ternate. Sedangkan sampelnya adalah seluruh populasi hal ini dikarenakan jumlah populasi yang sedikit sehingga seluruh populasi dijadikan sebagai sampel.

2.2. Metode Klasifikasi

Metode klasifikasi merupakan salah satu metode yang ada didalam machine learning. Didalam metode klasifikasi terdapat beberapa algoritma seperti Support Vector Machine dan Jaringan Saraf Tiruan sebagaimana yang akan diterapkan dalam penelitian ini. Cara kerjanya algoritma klasifikasi yaitu dengan data-data dikumpulkan kemudian dibagi kedalam data training dan data testing bahkan bisa ditambah lagi dengan data validasi. Selanjutnya data training dilakukan pelatihan sehingga terbentuk model agar data testing bisa dilakukan pengujian berdasarkan model yang telah terbentuk. Adapun Langkah-langkah dalam menerapkan metode klasifikasi ditunjukkan pada Gambar 2.



Gambar 2. Tahapan penerapan metode klasifikasi

1. Langkah awal dalam proses klasifikasi yaitu dataset yang telah dikumpulkan kemudian dibaca secara otomatis. Adapun dataset yang dibaca adalah data yang telah ditentukan label atau hasil keputusan dari pakar dalam ini dokter terkait status seorang pasien suspek TBC apakah yang bersangkutan beresiko atau tidak beresiko.

2. Langkah berikutnya adalah dilakukan rekayasa fitur, dimana akan dipilih fitur atau kriteria yang layak dan tidak layak dalam proses klasifikasi.
3. Selanjutnya akan diperiksa seluruh nilai fitur apakah semuanya sudah lengkap (L) atau ada yang tidak lengkap (TL) dalam hal ini disebut missing value
4. Jika masih terdapat missing value maka akan dilakukan proses imputasi. Selanjutnya dicek Kembali apakah sudah tidak terdapat missing value. Jika memang sudah tidak terdapat maka akan dilakukan proses klasifikasi.
5. Proses klasifikasi akan dilakukan Ketika data-data semuanya sudah lengkap nilai-nilainya.
6. Selanjutnya pada langkah terakhir adalah diuji kinerja algoritma dengan menggunakan beberapa metode semisal uji akurasi, presisi, recall dan F Score.

### 2.3. Proses Imputasi

Tindakan imputasi data perlu dilakukan, hal ini agar memastikan seluruh data tidak ada yang hilang atau biasanya dikenal dengan missing value. Tujuan dari tahapan ini yaitu agar memastikan seluruh data telah lengkap sehingga proses penerapan model disertai dengan pengujian model dapat dilakukan.

### 2.4. Support Vector Machine dan Jaringan Saraf Tiruan

Support Vector Machine adalah bagian dari algoritma klasifikasi dari machine learning yang mana dasarnya dari pembelajaran statistik. Ide dasar dari SVM adalah memaksimalkan batas hyperlane agar dapat hyperlane terbaik yang berfungsi untuk pemisah dua buah kelas dataset. Hyperlane atau batas keputusan pemisah terbaik antara kedua kelas dapat ditemukan dengan mengukur margin hyperlane tersebut dan mencari titik maksimalnya. Margin adalah jarak antara hyperlane tersebut dengan data terdekat dari masing-masing kelas. Data yang paling dekat ini disebut support vector.

Berbeda dengan Support Vector Machine. Untuk jaringan saraf tiruan cara kerjanya dengan mengadopsi sistem saraf manusia. Dimana pada Artificial Neural Network terdapat sebuah unit pemroses yang disebut neuron, fungsi aktivasi, bobot dan sejumlah vector masukan atau nilai inputan. Adapun pada jaringan saraf tiruan ada algoritma yang bisa dilakukan proses klasifikasi dan ada juga yang bisa melakukan clustering. Namun pada penelitian hanya digunakan jaringan saraf tiruan untuk proses klasifikasi yaitu jaringan saraf tiruan backpropagation.

### 2.5. Mengukur Kinerja Algoritma

Evaluasi suatu algoritma adalah salah satu tahapan paling penting yang harus dilakukan dalam implementasi metode klasifikasi machine learning, dengan melakukan evaluasi algoritma maka akan diketahui kinerja dari setiap algoritma. Untuk proses evaluasi bisa dengan menggunakan menghitung akurasi. Cara mendapat nilai akurasi yaitu dengan membandingkan hasil keputusan sistem dari data uji dengan data sebenarnya dari dokter. Untuk persamaannya ditunjukkan pada Persamaan 1.

$$akurasi = \frac{\sum \text{data uji yang prediksi benar}}{\sum \text{seluruh data uji}} \quad (1)$$

Selain itu bisa juga dengan menggunakan *confusion matrix*. Pada *confusion matrix* terdapat beberapa hal semacam *True Positive*, *True Negative*, *False Positive* dan *False Negative*. Beberapa hal ini dibutuhkan agar dapat mencari nilai recall, presisi dan F1 Score. Adapun matriks confusion ditunjukkan pada Tabel 1.

Tabel 1. Confusion Matrix

Kelas Prediksi	Kelas Sebenarnya	
	Beresiko	Tidak Beresiko
Beresiko	TP	FN
Tidak Beresiko	FP	TP

Dimana: TP adalah True Positive, FN adalah False Negative, FP adalah False Positive dan TP merupakan True Positive. Nilai presisi, recall dan F Score didapatkan berdasarkan Persamaan 2, Persamaan 3 dan Persamaan 4.

$$Presisi = \frac{TP}{TP+FP} \dots \dots \dots (2)$$

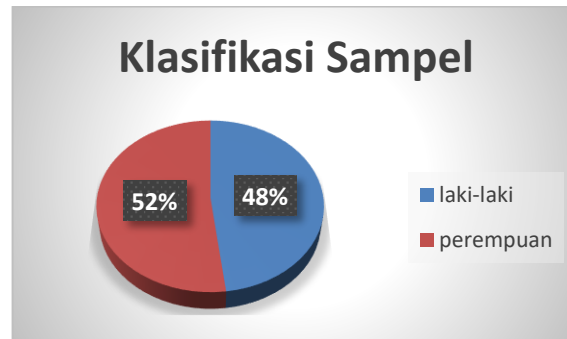
$$Recall = \frac{TP}{TP+FN} \dots \dots \dots (3)$$

$$F\ Score = 2 \cdot \frac{Presisi \cdot Recall}{Presisi + Recall} \dots \dots \dots (4)$$

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Populasi Dan Sampel

Populasi dari penelitian yaitu masyarakat yang berada di wilayah kerja Puskesmas Kalumata Kota Ternate. Untuk pengambilan sampel digunakan Teknik random sampling. Dengan jumlah sampel dari penelitian yaitu sebesar 100 sampe. Dari sampel yang didapat sampel terbesar dari kalangan perempuan yaitu sebanyak 52%, sebagaimana ditunjukkan pada Gambar 3.



Gambar 3. Klasifikasi sampel penelitian berdasarkan jenis kelamin

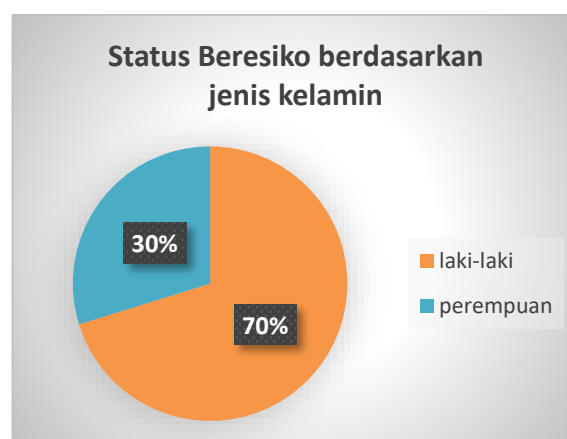
#### 3.2 Hasil Pelacakan

Data-data yang telah dikumpulkan kemudian dilakukan proses analisis statistic sederhana sebelum diterapkan metode klasifikasi. hasil pelacakan atau survey pada pasien suspek tubercolis didapat hasil 47% Beresiko sedangkan sisanya 53% tidak beresiko. Adapun gambarnya ditunjukkan pada Gambar 4.



Gambar 4. Hasil Pelacakan pasien suspek TBC

Hasil pelacakan ini juga menunjukkan Sebagian besar pasien yang beresiko Tuberculosis yaitu pasien dengan berjenis Kelamin Laki-laki. Dengan usia diatas 17 tahun. Adapun hasil perbandingan pasien yang beresiko berdasarkan jenis kelamin ditunjukkan pada Gambar 5.



Gambar 5. Status Resiko berdasarkan jenis kelamin

#### 3.3 Hasil Penerapan Metode Klasifikasi

Data yang telah dianalisis oleh dokter maka selanjutnya dilakukan penerapan metode klasifikasi dengan menggunakan langkah-langkah sebagaimana yang ditunjukkan pada Gambar 2. Adapun jumlah dataset dalam penelitian ini sebanyak 100 data. Sebagaimana ditunjukkan pada Gambar 6.

no	nama	jenis kelamin	usia	tingkat pendidikan	pekerjaan	alamat	pendapatan bulanan	Pernah TBC	keluarga anda TBC	...
0	1	AN	Laki-Laki	42.0	SMA	Wiraswasta	Bastiong	>1.000.000	tidak	tidak ...
1	2	Aulia	Perempuan	17.0	SMP	tidak bekerja	Bastiong	0	tidak	tidak ...
2	3	bapak anto	Laki-Laki	35.0	sma	Wiraswasta	tabona	>1.000.000	tidak	tidak ...
3	4	ibu sarina	Perempuan	43.0	sma	Wiraswasta	kayu merah	0	tidak	ya ...
4	5	fitria	Perempuan	33.0	S1	Wiraswasta	kayu merah	>1.000.000	tidak	ya ...

5 rows x 21 columns

Gambar 6. Hasil membaca dataset

Selanjutnya dilakukan proses penentuan kriteria berdasarkan data yang telah terbaca, sehingga beberapa kolom seperti nama pasien, tingkat pendidikan, alamat dan pendapatan bulanan tidak dimasukan dalam proses klasifikasi. Sehingga jumlah kriteria dalam penelitian yaitu sebanyak 13 kriteria. Setelah itu diperiksa nilai fitur dan didapatkan beberapa nilai pada kriteria yang masih terjadi *missing value* sebagaimana ditunjukkan pada Gambar 7.

```
In [45]: data.isnull().sum()[0:13]
Out[45]: jenis kelamin          0
         Pernah TBC             1
         Anggota keluarga anda TBC 0
         anda membuang dahak/luda disembarang tempat 0
         anda membuka jendela pada pagi dan siang hari 1
         tutup mulut saat batuk dan bersin 1
         merokok setiap hari 2
         anggota rumah tangga ada yang perokok 0
         mencuci tangan sebelum dan sesudah kegiatan 0
         makan makanan yang bergizi setiap hari 0
         merasakan gejala batuk lebih dari tiga minggu, demam disertai infulensa 0
         merasakan nyeri dada, sesak dada dan batuk berdarah 0
         menggunakan masker jika sedang dalam perjalanan 0
         dtype: int64
```

Gambar 7. Hasil pengecekan *missing value*

Data-data yang masih terdapat *missing value* tentu akan berpengaruh pada proses klasifikasi sehingga untuk penanganan *missing value* dilakukan proses imputasi. Dalam penelitian ini dilakukan teknik imputasi modus. Teknik imputasi modus cara kerjanya dengan mengganti semua kemunculan nilai yang hilang dalam suatu variabel dengan mode, atau nilai yang paling sering muncul dalam kriteria tersebut. Adapun hasil imputasi ditunjukkan pada Gambar 8.

```
In [49]: data.isnull().sum()[0:13]
Out[49]: jenis kelamin          0
         Pernah TBC             0
         Anggota keluarga anda TBC 0
         anda membuang dahak/luda disembarang tempat 0
         anda membuka jendela pada pagi dan siang hari 0
         tutup mulut saat batuk dan bersin 0
         merokok setiap hari 0
         anggota rumah tangga ada yang perokok 0
         mencuci tangan sebelum dan sesudah kegiatan 0
         makan makanan yang bergizi setiap hari 0
         merasakan gejala batuk lebih dari tiga minggu, demam disertai infulensa 0
         merasakan nyeri dada, sesak dada dan batuk berdarah 0
         menggunakan masker jika sedang dalam perjalanan 0
         dtype: int64
```

Gambar 8. Hasil proses imputasi

### 3. 4 Hasil Pengujian dengan Model SVM

Pengujian model dilakukan ketika seluruh data sudah tidak terdapat *missing value*. Dalam pengujian dengan metode SVM dilakukan beberapa kali pengujian yaitu dengan mengganti beberapa jumlah data training dan data testing, penggunaan kernel.



Tabel 2. Hasil pengujian algoritma SVM

Testing	Data Training	Data Testing	Kernel	Akurasi
1	50	50	Gaussian	92%
2	40	60	Gaussian	91,66%
3	60	40	Gaussian	92,5%
4	50	50	Linear	84%
5	40	60	Linear	85%
6	60	40	Linear	85%
7	50	50	Poly	88%
8	40	60	Poly	90%
9	60	40	Poly	90%
10	50	50	Sigmoid	52%

Hasil pengujian implementasi SVM yang ditunjukkan pada Tabel 2 terdapat beberapa perbedaan yang signifikan setidaknya akurasi paling rendah ketika digunakan kernel sigmoid yaitu 52% sedangkan akurasi tertinggi ketika digunakan kernel gaussian yaitu 92%. Untuk itu pada proses evaluasi algoritma untuk mencari nilai recall, presisi dan F1 Score digunakan kernel Gaussian dengan jumlah data training 50 dan data testing 50. Adapun hasilnya ditunjukkan pada Gambar 9.

```
In [389]: print(classification_report(ytest, hasilPrediksi))
```

	precision	recall	f1-score	support
0	0.88	0.96	0.92	24
1	0.96	0.88	0.92	26
accuracy			0.92	50
macro avg	0.92	0.92	0.92	50
weighted avg	0.92	0.92	0.92	50

Gambar 9. hasil evaluasi algoritma

Hasil evaluasi algoritma menunjukkan bahwa nilai presisi untuk kategori beresiko lebih kecil yaitu 88% dibandingkan dengan nilai presisi tidak beresiko 96%. Sedangkan untuk nilai recall kategori beresiko nilai recall lebih tinggi 96% dibandingkan dengan nilai recall tidak beresiko 88%. Adapun untuk nilai F Score 92%.

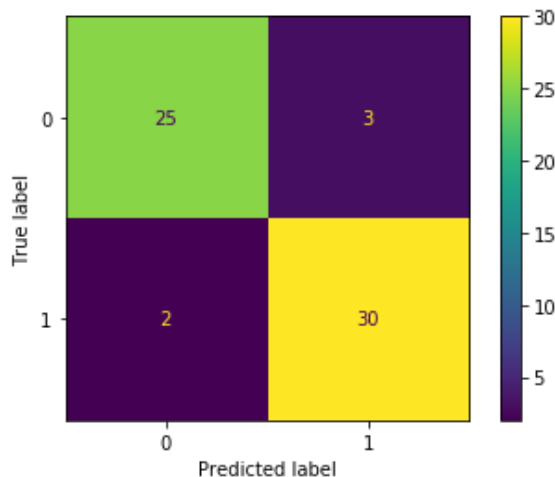
### 3. 5 Hasil pengujian dengan model JST

Dalam model Jaringan Saraf Tiruan (JST) terdapat beragam macam algoritma, namun pada penelitian digunakan *multi layer perceptron*. Dalam pengujian dengan JST juga dilakukan beberapa kali dengan mengkomparasi parameter tuning dari JST seperti jumlah iterasi, fungsi aktivasi, jumlah hidden layer dan menginisialisasi jumlah data testing dan data training yang berbeda-beda. Adapun hasil dari pengujian ditunjukkan pada Tabel 3.

Tabel 3. hasil pengujian dengan model JST

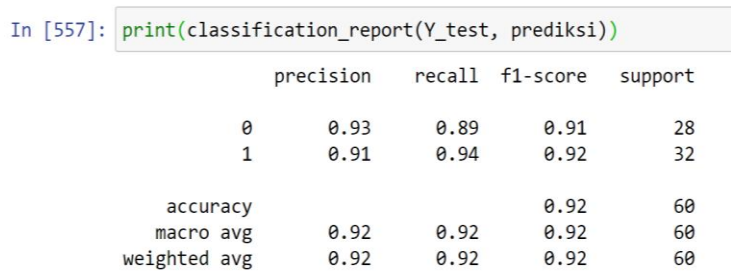
Dataset	Jumlah iterasi	Aktivasi	Validasi training	Akurasi
50:50	10000	Relu	80 %	88% %
40:60	10000	Relu	83,33%	91,66%
60:40	10000	Relu	87,5%	87,5%
50:50	5000	Relu	84%	84%
40:60	5000	Relu	79,16	90%
60:40	5000	Relu	83,33%	85%

Hasil pengujian menunjukkan bahwa jumlah literasi sangat berpengaruh terhadap hasil kinerja algoritma sebagaimana yang ditunjukkan pada Tabel 3. terdapat perbedaan antara inisialisasi iterasi antara 5000 dan 10000. akurasi tertinggi terdapat pada pengujian kedua dimana data training yang digunakan 40 data dan data testing 60 data dengan tingkat akurasi sebesar 91,66%. Adapun gambar hasil pengujian kedua ditunjukkan pada Gambar 10.



Gambar 10. Hasil pengujian kedua JST

Selanjutnya dilakukan evaluasi *recall*, presisi dan F1 Score untuk pengujian dengan Jaringan Saraf Tiruan Multi Layer Perceptron didapat nilai presisi untuk beresiko sebesar 93% dan tidak beresiko sebesar 91% dengan rata-rata secara keseluruhan nilai presisi 92%. untuk recall didapat nilai beresiko sebesar 89% dan tidak beresiko 94% dengan rata-rata nilai secara keseluruhan untuk recall sebesar 91,5 %. Adapun hasil evaluasi algoritma ditunjukkan pada Gambar 11.



Gambar 10. Hasil Evaluasi algoritma JST

### 3. 6 Hasil Penerapan K-Fold Cross Validation

Hasil akurasi yang telah didapatkan dari setiap algoritma kemudian dilakukan proses validasi dengan menerapkan k-fold cross validasi. Tujuan menggunakan k-fold cross validasi agar seluruh data yang terapkan memiliki hak yang sama yaitu sebagai data training dan sebagai data uji. Adapun hasil dari k-fold cross validasi untuk metode SVM didapat rata-rata akurasi untuk 2-fold yaitu 82%, 3 fold 83,16 %, 4 fold 83% dan 5 fold 82%. Sedangkan ketika menerapkan jaringa saraf tiruan untuk proses validasi dengan k-fold cross validasi didapatkan rata-rata akurasi untuk 2-fold 82%, 3-fold 85,08%, 4-fold 85% dan 5-fold 82,01%. dengan demikian jika menerapkan proses validasi akurasi tertinggi ada pada metode jaringan saraf tiruan yaitu ketika dibagi menjadi 3-Fold dimana akurasi yang didapatkan sebesar 85.08%.

## 4. KESIMPULAN

Berdasarkan hasil pelacakan pasien suspek TBC diwilayah kerja Puskesmas Kalumata kota Ternate untuk mengetahui apakah beresiko dan tidak beresiko dapat disimpulkan bahwa :

1. Dari 100 sampel yang didapatkan 47% beresiko untuk terjangkit tuberkolosis dimana dari 47% yang beresiko sebanyak 70% berjenis kelamin laki-laki.
2. Selanjutnya dilakukan proses penerapan algoritma klasifikasi Support Vector Machine dan Jaringan Saraf Tiruan akurasi tertinggi terdapat pada metode Support Vector Machine yaitu sebesar 92,5% sedangkan pada Jaringan Saraf Tiruan 91,66 %.
3. Ketika diterapkan k-fold cross validasi rata-rata akurasi tertinggi ketika diterapkan metode jaringan saraf tiruan dengan rata-rata akurasi sebesar 85,08% dengan 3-fold.



**DAFTAR PUSTAKA**

- [1] ANDAYANI, S. DAN ASTUTI, Y. (2017) "Prediksi Kejadian Penyakit Tuberkulosis Paru Berdasarkan Usia di Kabupaten Ponorogo Tahun 2016-2020", *Indonesian Journal For Health Sciences*, 1(2), pp. 29-33. 10.24269/ijhs.v1i2.482.
- [2] WHO. 2020. Global Tuberculosis Report., 99–117. <https://apps.who.int/iris/bitstream/handle/10665/336069/9789240013131-eng.pdf>, diakses tanggal 7 Juni 2021
- [3] Kemenkes., 2020., Pusat Data dan Informasi Kemenkes RI. 2020. "Info DATIN Tuberkulosis." <https://pusdatin.kemkes.go.id/article/view/18101500001/infodatin-tuberkulosis-2018.html> diakses tanggal 7 Juni 2021
- [4] Dinkes. 2020., Dinas Kesehatan: Profil Kesehatan Kota Ternate." Available : Rri.Co.Id/Ternate/Daerah. Retrieved June 19, 2021 (<https://rri.co.id/ternate/daerah/686339/dinkes-dan-puskesmas-se-kota-ternate-siap-layani-keluhan-warga>). diakses 19 Juni 2021.
- [5] Badeji. B., dan Ajisafe., 2018., Bayesian Classification Model in Predicting Tuberculosis Infection. *IOSR Journal of Computer Engineering (IOSR-JCE)*, vol 20 (4), pp. 6-16. doi: 10.9790/0661-2004010616
- [6] Amrin dan Saiyar, H., 2018., Aplikasi Diagnosa Penyakit Tuberculosis Menggunakan Algoritma Naive Bayes, *Jurnal Riset Komputer (JURIKOM)*, vol 5(5), pp. 498-502.
- [7] Aguiar. F. S., dkk., 2012., " Classification and regression tree (CART) model to predict pulmonary tuberculosis in hospitalized patients", *BMC Pulmonary Medicine*. available at : <http://www.biomedcentral.com/1471-2466/12/40>
- [8] Lumbanraja. F. R., 2020., Prediksi Jumlah Penderita Penyakit Tuberculosis Di Kota Bandar Lampung Menggunakan Metode Svm (Support Vector Machine)., *Kumpulan jurnaL Ilmu Komputer (KLIK)*, vol 7 (3), pp. 320-330.
- [9] Anshori, M., dkk., 2019., Preprocessing Approach for Tuberculosis DNA Classification using Support Vector Machines (SVM)., *Journal of Information Technology and Computer Science*. vol 4 (3), pp. 233-240.
- [10] Depinta. L., dan Abdullah. Z., 2017. Implementasi Jaringan Syaraf Tiruan Backpropagation untuk Deteksi Penyakit Tuberculosis (TB) Paru dari Citra Rontgen, *Jurnal Fisika Unand*. vol 6 (1), pp. 61-66
- [11] Ramadan, M, Y., Syaui, D., dan Tibyani., 2019., Implementasi Metode Klasifikasi Support Vector Machine (SVM) Terhadap Pemakaian Minyak Goreng., *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*. vol 3 (2), pp.1669-1677.
- [12] Yanti, N., dkk., 2018., Jaringan Syaraf Tiruan Untuk Pengenalan Citra Sidik Jari Pada Smart Home., *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*. 5 (5), pp. 597-604.
- [13] Tempola, F., Muhammad, M., dan Khairan, A., 2018., Perbandingan Klasifikasi Antara KNN dan Naive Bayes pada Penentuan Status Gunung Berapi dengan K-Fold Cross Validation., *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)* 5 (5), Pp. 577-584.
- [14] Kumar, K. V., dkk., 2020., Performance Evaluation Of Machine Learning Algorithms For Disease Prediction., *International Journal of Advanced Science and Technology*, vol 29 (7), pp. 7820-7830
- [15] Surbakti, A. Q., Hayami, R., dan Amien, J. A., dkk., 2021., Analisa Tanggapan Terhadap Psbb Di Indonesia Dengan Algoritma Decision Tree Pada Twitter., *Jurnal Computer Science and Information Technology (CoSciTech)*, vol 2 (2), pp. 91-97
- [16] Mukhtar, H., Rifaldo, M., Taufiq, R. M., dan Rizki, Y., dkk., 2021., Peramalan Kedatangan Wisatawan Mancanegara Ke Indonesia Menurut Kebangsaan Perbulannya Menggunakan Metode Multilayer Perceptron, *Jurnal Computer Science and Information Technology (CoSciTech)*, vol 2 (2), pp. 113-119