

**DETEKSI PENYAKIT DIABETES IBU HAMIL MENGGUNAKAN  
METODE *ARTIFICIAL NEURAL NETWORK***

**SKRIPSI**

**Oleh:  
WINDYA ZAHRAH KAMALI  
NIM. 19650150**



**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2023**

**DETEKSI DIABETES IBU HAMIL MENGGUNAKAN  
METODE ARTIFICIAL NEURAL NETWORK**

**SKRIPSI**

**Oleh :  
WINDYA ZAHRAH KAMALI  
NIM. 19650150**

Diajukan kepada:  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri (UIN) Malang  
Untuk memenuhi Salah Satu Persyaratan dalam  
Memperoleh Gelar Sarjana Komputer (S. Kom)

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2023**

**HALAMAN PERSETUJUAN**

**DETEKSI PENYAKIT DIABETES IBU HAMIL MENGGUNAKAN  
METODE *ARTIFICIAL NEURAL NETWORK***

**SKRIPSI**

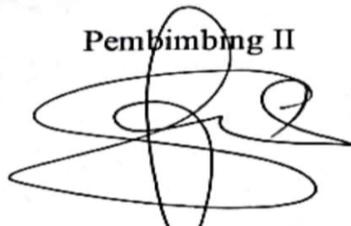
**Oleh:  
WINDYA ZAHRAH KAMALI  
NIM. 19650150**

**Telah Diperiksa dan Disetujui untuk Diuji  
Tanggal:**

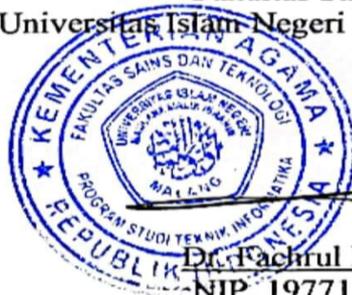
**Pembimbing I**

  
Dr. Cahyo Crysdiyan  
NIP. 1974024 2009011 1 008

**Pembimbing II**

  
Dr. M. Amin Haryadi M.T  
NIP. 19670018 200501 1 001

**Mengetahui,  
Ketua Program Studi Teknik Informatika  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang**



  
Dr. Fachrul Kurniawan, M.MT  
NIP. 19771020 200912 1 001

## HALAMAN PENGESAHAN

### DETEKSI PENYAKIT DIABETES IBU HAMIL MENGGUNAKAN METODE *ARTIFICIAL NEURAL NETWORK*

#### SKRIPSI

Oleh:  
**WINDYA ZAHRAH KAMALI**  
NIM. 19650150

Telah Dipertahankan di Depan Dewan Penguji  
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan  
untuk Memperoleh Gelar Sarjana Komputer (S.Kom)  
Pada Tanggal: 13 Desember 2023

#### Susunan Dewan Penguji

Ketua Penguji : Dr. Ririen Kusumawati, S.Si., M.Kom  
NIP. 19720309 200501 2 002

Anggota Penguji I : Dr. Irwan Budi Santoso, M.Kom  
NIP. 19770103 201101 1 004

Anggota Penguji II : Dr. Cahyo Crysdiان  
NIP. 1974024 2009011 1 008

Anggota Penguji III : Dr. M. Amin Haryadi M.T  
NIP. 19670018 200501 1 001

( )  
( )  
( )  
( )

Mengetahui dan Mengesahkan,  
Ketua Program Studi Teknik Informatika  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Dr. Fachrul Kurniawan, M.MT  
NIP. 19771020 200912 1 001

## PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Windya Zahrah Kamali  
NIM : 19650150  
Fakultas/Prodi : Sains dan Teknologi / Teknik Informatika  
Judul Skripsi : Deteksi Diabetes Ibu Hamil Menggunakan Metode *Artificial Neural Network*

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 13 Desember 2023  
Yang membuat pernyataan,



Windya Zahrah Kamali  
NIM. 1965003

**HALAMAN MOTTO**

**“It’s not always easy, but that’s life, be strong because there are better days ahead”**

-Mark Lee From NCT

## **HALAMAN PERSEMBAHAN**

Puji Syukur atas kehadiran Allah *Subhanahu wa ta'ala*  
Shalawat serta salam kepada Rasulullah *Shallallahu 'alaihi wasallam*

Penulis mempersembahkan karya ini kepada kedua orang tua, dosen, kakak, sahabat, dan semua pihak yang telah memberikan motivasi, doa, dan dukungan sehingga skripsi ini dapat selesai pada waktunya.

## KATA PENGANTAR

*Assalamualaikum Wr. Wb*

Puji dan syukur penulis panjatkan kehadiran Tuhan Yang Maha Esa, Allah subhanahu wa ta'ala yang telah memberikan Taufik dan Hidayah-Nya kepada penulis sehingga dapat menyelesaikan skripsi yang berjudul “Deteksi Diabetes Ibu Hamil Menggunakan Metode *Artificial Neural Network*” dengan baik.

Banyak pihak yang terlibat dalam penulisan skripsi ini yang telah memberikan dukungan baik moril maupun materil. Untuk itu dalam kesempatan kali ini penulis ingin mengucapkan banyak terimakasih kepada:

1. Prof. Dr. H.M. Zainuddin, MA selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim
2. Prof. Dr. Sri Hariani, M.Si selalu dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim
3. Dr. Fachrul Kurniawan, M.MT selaku Ketua Jurusan Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang yang senantiasa memberikan dorongan untuk menyelesaikan skripsi ini.
4. Dr. Cahyo Crysdiyan selaku Dosen Pembimbing I yang telah dengan sabar memberikan arahan baik dalam penulisan hingga program yang dibuat dalam menyelesaikan skripsi ini.
5. Dr. M. Amin Hariyadi M.T selaku Dosen Pembimbing II yang telah meluangkan waktu untuk memberikan ilmu dan arahan dalam menyelesaikan skripsi ini.
6. Seluruh dosen dan staff akademika Jurusan Teknik Informatika yang memberikan ilmu yang sangat bermanfaat yang secara tidak langsung ikut terlibat dalam penyusunan skripsi ini.
7. Kepada kedua orang tua saya yang saya cintai, Bapak Musthofa dan Ibu Luluk yang telah memberikan suntikan dana terbesar dan selalu memberikan doa, semangat serta dukungan untuk menyelesaikan perkuliahan dan skripsi ini.

Serta kakak saya Hendra, Akbar, Dewi, Selly yang telah memberikan motivasi untuk terus menyelesaikan skripsi ini.

8. Kepada mbak Nia Faricha S.Si tersayang terima kasih telah membantu menyusun jadwal hingga diusahakan bisa sampai reschedule.
9. Seluruh keluarga besar Saudara Teknik Informatika UIN Maulana Malik Ibrahim Malang terkhusus Angkatan 2019 *Alliance Of Informatics "Alien"* serta Grup YG Swag yang beranggotakan "Bila, Endah, dan Mauri" atas segala rangkaian cerita, dukungan bantuan dalam bentuk apapun selama perkuliahan hingga terselesaikannya penyusunan skripsi ini.
10. Seluruh pihak yang telah terlibat secara langsung maupun tidak langsung dalam proses penyusunan skripsi sejauh ini.
11. *Last but not least. I wanna thank me, i wanna thank me for believing in me, i wanna thank me for doing all this hard work, i wanna thank me for having no days off, i wanna thank me for never quitting, i wanna thank me for just being me of all time.*

Penulis menyadari dalam penulisan skripsi ini tidak luput dari kesalahan yang jauh dari kata sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun sehingga skripsi ini dapat lebih dikembangkan dan berguna bagi penulis sendiri maupun pembaca pada umumnya.

Malang, 13 Desember 2023

Penulis

## DAFTAR ISI

<b>HALAMAN PERSETUJUAN</b> .....	Kesalahan! Bookmark tidak ditentukan.
<b>HALAMAN PENGESAHAN</b> .....	<b>iii</b>
<b>PERNYATAAN KEASLIAN TULISAN</b> .....	Kesalahan! Bookmark tidak ditentukan.
<b>HALAMAN MOTTO</b> .....	<b>vi</b>
<b>HALAMAN PERSEMBAHAN</b> .....	<b>vii</b>
<b>KATA PENGANTAR</b> .....	<b>viii</b>
<b>DAFTAR ISI</b> .....	<b>x</b>
<b>DAFTAR GAMBAR</b> .....	<b>xi</b>
<b>DAFTAR TABEL</b> .....	<b>xii</b>
<b>ABSTRAK</b> .....	<b>xiii</b>
<b>ABSTRACT</b> .....	<b>xiv</b>
مستخلص البحث .....	<b>xv</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Pernyataan Masalah .....	4
1.3 Tujuan Penelitian .....	5
1.4 Batasan Masalah .....	5
1.5 Manfaat Penelitian .....	5
<b>BAB II STUDI PUSTAKA</b> .....	<b>7</b>
2.1 Deteksi Penyakit Diabetes .....	7
2.2 <i>Artificial Neural Network</i> dalam Mendeteksi Penyakit .....	12
<b>BAB III METODOLOGI PENELITIAN</b> .....	<b>18</b>
3.1 Pengumpulan Data .....	18
3.2 Desain Sistem.....	19
3.3 <i>Artificial Neural Network</i> .....	21
<b>BAB IV UJI COBA DAN PEMBAHASAN</b> .....	<b>33</b>
4.1 Skenario Uji Coba .....	33
4.1.1 Data Penelitian .....	33
4.1.2 Menghitung Kinerja Sistem .....	35
4.2 Hasil Uji Coba.....	37
4.2.1 Uji Coba Ratio Data 70:30.....	37
4.2.2 Uji Coba Ratio Data 80:20.....	39
4.2.3 Hasil Uji Coba Ratio 90:10.....	41
4.3 Pembahasan.....	44
<b>BAB V PENUTUP</b> .....	<b>49</b>
5.1 Kesimpulan .....	49
5.2 Saran.....	49
<b>DAFTAR PUSTAKA</b>	
<b>LAMPIRAN</b>	

## DAFTAR GAMBAR

Gambar 3.1 Desain Sistem.....	20
Gambar 3.2 Arsitektur Artificial Neural Network (Pohan <i>et al.</i> 2020) .....	22
Gambar 3.3 <i>Flowchart Backpropagation</i> <b>Kesalahan! Bookmark tidak ditentukan.</b>	
Gambar 3.4 Alur Proses <i>Training</i> .....	30
Gambar 3.5 Alur Proses <i>Testing</i> .....	32
Gambar 4.1 Source Code Normalisasi .....	34
Gambar 4.2 Source Code Split Data .....	35
Gambar 4.3 Confusion Matrix 70:30 .....	37
Gambar 4.4 Confusion Matrix 70:30 pada learning rate 0.08 .....	38
Gambar 4.5 <i>Confusion Matrix</i> Hasil Deteksi.....	40
Gambar 4.6 Confusion Matrix 80:20 pada learning rate 0.08 .....	41
Gambar 4.7 Confusion Matrix 90:10, learning rate 0.04 .....	42
Gambar 4.8 Confusion Matrix 90:10, <i>learning rate</i> 0.08 .....	43

## DAFTAR TABEL

Tabel 3.1 Fitur Dataset.....	18
Tabel 3.2 Contoh Dataset Penyakit Diabetes.....	19
Tabel 3.3 Contoh data pasien sebelum dilakukan normalisasi .....	21
Tabel 4.1 Hasil Normalisasi.....	34
Tabel 4.3 Confusion Matrix .....	36
Tabel 4.4 Nilai <i>accuracy, precision, recall</i> ratio 70:30, <i>learning rate</i> 0.04 ..	38
Tabel 4.5 Nilai <i>accuracy, precision, recall</i> ratio 70:30, <i>learning rate</i> 0.08 ....	39
Tabel 4.6 Nilai <i>accuracy, precision, recall</i> .....	40
Tabel 4. 7 Nilai <i>accuracy, precision, recall</i> .....	41
Tabel 4.8 Nilai <i>accuracy, precision, recall</i> .....	42
Tabel 4.9 Nilai <i>Accuracy, Precision, Recall</i> 90:10, <i>learning rate</i> 0.08 .....	43
Tabel 4. 10 Hasil Akurasi Seluruh Ratio .....	43

## ABSTRAK

Kamali, Windya Zahrah. 2023. **Deteksi Penyakit Diabetes Ibu Hamil Menggunakan Metode *Artificial Neural Network***. Skripsi, Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Cahyo Crysdiyan. (II) Dr. M. Amin Hariyadi.

**Kata Kunci:** *Artificial Neural Network, Backpropagation, Confusion Matrix, Diabetes, Penyakit.*

Diabetes merupakan salah satu penyakit serius dan fatal di Indonesia, bersanding dengan stroke dan jantung koroner. Penyakit ini menjadi penyebab utama masalah kardiovaskular dan berkontribusi pada tingginya angka kematian ibu saat melahirkan. Selain itu, diabetes juga dapat menularkan risiko penyakit kepada bayi yang baru lahir. Dalam menghadapi kondisi ini, deteksi dini diabetes menjadi krusial untuk mencegah komplikasi yang lebih serius. Pemanfaatan teknologi dianggap sebagai solusi untuk mengurangi kesalahan estimasi. Dalam penelitian ini, digunakan metode Jaringan Saraf Tiruan (*Artificial Neural Network*) untuk mendeteksi diabetes pada ibu hamil. Metode ini dikenal karena kemampuannya dalam mengenali pola-pola kompleks dan dapat diaplikasikan pada berbagai jenis data. Dataset yang digunakan berasal dari Institut Nasional Diabetes dan Pencernaan serta Penyakit Ginjal, terdiri dari 768 data, dan uji coba dilakukan dengan menggunakan metode *Confusion Matrix*. Penelitian ini bertujuan untuk menentukan tingkat akurasi dari sistem yang dikembangkan. Dataset dibagi menjadi dua bagian, dan dilakukan eksperimen dengan mengatur parameter seperti *learning rate*, jumlah *hidden layers*, dan *epochs*. Hasil penelitian menunjukkan bahwa akurasi tertinggi tercapai pada penggunaan model 3, yaitu menggunakan 90% dari jumlah data, *learning rate* sebesar 0,04, 50 *epochs*, dengan nilai akurasi mencapai 80.5%, presisi sebesar 83.3%, dan *recall* sebesar 64.5%. Dari hasil uji coba, dapat disimpulkan bahwa tingkat akurasi meningkat seiring dengan peningkatan jumlah data dan *epochs*, sejalan dengan jumlah *hidden layers*, serta besarnya *learning rate* yang digunakan.

## ABSTRACT

Kamali, Windya Zahrah. 2023. **Deteksi Penyakit Diabetes Ibu Hamil Menggunakan Metode *Artificial Neural Network***. Undergraduate Thesis, Department of Informatics Engineering, Faculty of Science and Technology, Maulana Malik Ibrahim State Islamic University Malang. Advisors: (I) Dr. Cahyo Crysdiyan. (II) Dr. M. Amin Hariyadi.

Diabetes is one of the most serious and fatal diseases in Indonesia, along with stroke and coronary heart disease. It is the leading cause of cardiovascular problems and contributes to the high maternal mortality rate during childbirth. In addition, diabetes can also transmit the risk of disease to newborn babies. In the face of this condition, early detection of diabetes is crucial to prevent more serious complications. The utilization of technology is considered as a solution to reduce estimation errors. In this study, the Artificial Neural Network method is used to detect diabetes in pregnant women. This method is known for its ability to recognize complex patterns and can be applied to various types of data. The dataset used came from the National Institute of Diabetes and Digestive and Kidney Diseases, consisting of 768 data, and testing was carried out using the Confusion Matrix method. This study aims to determine the accuracy of the developed system. The dataset was divided into two parts, and experiments were conducted by setting parameters such as learning rate, number of hidden layers, and epochs. The results showed that the highest accuracy was achieved using model 3, which used 90% of the total data, learning rate of 0.04, 50 epochs, with an accuracy value of 80.5%, precision of 83.3%, and recall of 64.5%.

**Keywords:** *Artificial Neural Network, Backpropagation, Confusion Matrix, Diabetes, Disease.*

## مستخلص البحث

كمالي، وينديا زهرة. ٢٠٢٣. الكشف عن مرض السكري لدى النساء الحوامل باستخدام طريقة الشبكة العصبية الاصطناعية. البحث الجامعي، قسم الهندسة المعلوماتية، كلية العلوم والتكنولوجيا بجامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرف الأول: د. جاهيو كريسديان. المشرف الثاني: د. أمين هارياي.

**الكلمات الرئيسية:** الشبكة العصبية الاصطناعية، الانتشار الخلفي، مصفوفة الارتباك، مرض السكري، المرض.

مرض السكري هو أحد الأمراض الخطيرة والمميتة في إندونيسيا، إلى جانب السكتة الدماغية وأمراض القلب التاجية. هذا المرض هو سبب رئيسي لمشاكل القلب والأوعية الدموية ويساهم في ارتفاع معدل وفيات الأمهات أثناء الولادة. بالإضافة إلى ذلك، يمكن أن ينقل مرض السكري أيضا خطر الإصابة بالأمراض إلى الأطفال حديثي الولادة. في التعامل مع هذه الحالة، يعد الاكتشاف المبكر لمرض السكري أمرا بالغ الأهمية لمنع حدوث مضاعفات أكثر خطورة. يعتبر استخدام التكنولوجيا حلا لتقليل أخطاء التقدير. في هذا البحث، تم استخدام طريقة الشبكة العصبية الاصطناعية للكشف عن مرض السكري لدى النساء الحوامل. تعرف هذه الطريقة بقدرتها على التعرف على الأنماط المعقدة ويمكن تطبيقها على العديد من أنواع البيانات المختلفة. جاءت مجموعة البيانات المستخدمة من المعهد الوطني للسكري وأمراض الجهاز الهضمي والكلى، وتألقت من ٧٦٨ بيانات، وأجريت التجربة باستخدام طريقة مصفوفة الارتباك. يهدف هذا البحث إلى معرفة مستوى دقة النظام المتطور. تم تقسيم مجموعة البيانات إلى جزأين، وتم إجراء التجارب من خلال تحديد معلمات مثل معدل التعلم وعدد الطبقات المخفية والعصور. أظهرت النتائج أنه تم تحقيق أعلى دقة في استخدام النموذج ٣، الذي استخدم ٩٠% من كمية البيانات، ومعدل التعلم ٠.٠٤، و ٥٠ عصارا، بقيمة دقة ٨٠.٥%، وثبات ٨٣.٣%، واستدعاء 64.5%. من نتائج هذه التجربة، يمكن الاستنتاج أن مستوى الدقة يزداد مع زيادة كمية البيانات والعصور، بما يتماشى مع عدد الطبقات المخفية، فضلا عن معدل التعلم الكبير المستخدم.

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Diabetes adalah penyakit yang menyebabkan kebutaan, gagal ginjal dan serangan jantung dan kematian. Menurut International Diabetes Federation (IDF), 463 juta orang di seluruh dunia menderita diabetes pada tahun 2019. Jumlah ini diperkirakan akan meningkat sebesar 578 juta pada tahun 2030 dan menjadi 700 juta pada tahun 2045. Indonesia sendiri merupakan salah satu dari 10 negara dengan jumlah penderita diabetes terbanyak. diabetes di dunia tahun 2019 (Carracher *et al.*, 2018).

Menentukan jenis penyakit diabetes memerlukan keterampilan seorang pakar atau pakar (Riadi, 2017). Banyak orang yang terlambat mengetahui kondisinya, sehingga saat diperiksa sudah dalam kondisi serius (Hadi *et al.*, 2018). Menanggapi situasi ini, deteksi dini diabetes diperlukan untuk mencegah penyakit yang lebih kompleks. Penggunaan teknologi dapat digunakan untuk meminimalkan kesalahan estimasi. Dalam dunia kedokteran, teknologi memungkinkan para ahli atau dokter untuk menginterpretasikan penyakit dengan cepat dan akurat (Gunawan *et al.*, 2020).

Teknologi yang dapat digunakan untuk mengekstraksi informasi dari data yang ada disebut data mining (Budiyasari *et al.*, 2017). Data mining dapat digunakan untuk memprediksi kejadian di masa depan (Prajarini *et al.*, 2016). Salah satunya memprediksi diabetes. Dalam data mining terdapat teknik klasifikasi, yaitu

suatu metode analisis dengan memprediksi identifier atau kelas dari sampel data untuk menemukan model yang tepat untuk dapat mengklasifikasikan data baru yang akan dianalisis (Purwanto & Darmadi, 2018).

Proses diagnosa penyakit bagi tenaga medis mempunyai peranan yang sangat penting, tidak hanya tingkat akurasi yang tinggi, tetapi juga tingkat kecepatan dalam mendiagnosis penyakit yang diderita pasien. Apabila tenaga medis mengalami kesalahan atau keterlambatan dalam mendiagnosis penyakit pasien, maka bentuk pengobatan dan penanganan terhadap pasien juga akan salah. Oleh sebab itu, dalam proses diagnosis penyakit terutama penyakit yang berhubungan dengan organ pernapasan, diperlukan diagnosa dengan tingkat akurasi yang tinggi dan proses diagnosis yang cepat (Amato *et al.*,2013).

Seperti yang dijelaskan Al-Qur'an Q.S Yusuf ayat 47-49 yang berbunyi (Kemenag, 2002):

قَالَ تَزْرَعُونَ سَبْعَ سِنِينَ دَابَّاتٍ فَمَا حَصَدْتُمْ فَذَرُوهُ فِي سُنْبُلِهِ إِلَّا قَلِيلًا مِّمَّا تَأْكُلُونَ (٤٧) ثُمَّ يَأْتِي مِنْ بَعْدِ ذَلِكَ سَبْعٌ شِدَادٌ يَأْكُلْنَ مَا قَدَّمْتُمْ لَهُنَّ إِلَّا قَلِيلًا مِّمَّا تَحْصِنُونَ (٤٨) ثُمَّ يَأْتِي مِنْ بَعْدِ ذَلِكَ عَامٌ فِيهِ يُغَاثُ النَّاسُ وَفِيهِ يَعْرِضُونَ (٤٩)

*“Dia (Yusuf) berkata, “Agar kamu bercocok tanam tujuh tahun (berturut-turut) sebagaimana biasa; kemudian apa yang kamu tuai hendaklah kamu biarkan di tangkainya kecuali sedikit untuk kamu makan. 47. Kemudian setelah itu akan datang tujuh (tahun) yang sangat sulit, yang menghabiskan apa yang kamu simpan untuk menghadapinya (tahun sulit), kecuali sedikit dari apa (bibit gandum) yang kamu simpan. 48. Setelah itu akan datang tahun, di mana manusia diberi hujan (dengan cukup) dan pada masa itu mereka memeras (anggur)” 49. (QS. Yusuf ayat 47-49).*

Ayat di atas menjelaskan bahwa Nabi Yusuf AS ketika menafsirkan mimpi raja meramalkan akan datangnya tujuh tahun kelaparan. Nabi Yusuf AS meminta kepada Raja untuk menyiapkan perbekalan selama tujuh tahun untuk

menghadapi musim ini berdasarkan wahyu dari Allah. Nubuatan Nabi Yusuf berfungsi untuk memberi tahu raja bahwa masa-masa sulit sedang mendekat dan memungkinkannya menyiapkan perbekalan yang cukup untuk menghadapi kesulitan yang akan datang. Paragraf ini sesuai dengan tujuan dan manfaat penelitian yang dilakukan oleh penulis, yaitu untuk mengetahui dan mendeteksi apakah seseorang mengidap penyakit diabetes sehingga penyebab kemungkinan terjadinya penyakit diabetes dapat dicegah dan dihindari.

Pada penelitian ini penyakit diabetes diprediksikan melalui beberapa faktor resiko yang tersedia di dalam data yang peneliti dapat dari kumpulan data ini berasal dari Institut Nasional Diabetes dan Pencernaan dan Penyakit Ginjal yang berada di situs resmi *Kaggle* yang mana dalam data tersebut berusia minimal 21 tahun. Prediksi ini menggunakan *Artificial Neural Network* yang akan memutuskan bahwa seorang tersebut terkena diabetes atau tidak. ANN salah satu metode yang penggunaannya telah lama diuji tingkat keakuratan data dengan beberapa fitur pelatihan jaringan.

Jaringan saraf tiruan adalah sistem pemrosesan informasi cerdas yang mewakili pengembangan model matematika umum. Prinsip kerja ANN terinspirasi dari prinsip kerja sistem jaringan saraf manusia. Para ilmuwan telah mengembangkan algoritma matematika yang bekerja serupa dengan pola perilaku neuron. Oleh karena itu digunakanlah nama ANN dan dalam bahasa Indonesia biasa disebut Artificial Neural Network (ANN). ANN adalah pendekatan yang berbeda dari metode AI lainnya. JST merupakan model kecerdasan yang terinspirasi dari struktur otak manusia dan diimplementasikan dengan menggunakan program

komputer yang dapat melakukan serangkaian perhitungan selama proses pembelajaran. JST terdiri dari sekumpulan prosesor sederhana yang saling berhubungan yang disebut neuron. Neuron yang dihubungkan oleh bobot meneruskan sinyal dari satu neuron ke neuron lainnya (Kurniawansyah, 2018).

Pada penelitian-penelitian sebelumnya, metode JST ini telah banyak digunakan dalam bidang medis untuk mendiagnosis penyakit seperti tuberkulosis dengan tingkat akurasi sebesar 95.08%, pneumonia dengan tingkat akurasi 82.08%, PPOK dengan tingkat akurasi 96.08%, dan asma akurasi 90%. Karena kemampuan generalisasinya, ANN dapat memecahkan masalah sistem yang kompleks seperti klasifikasi gambar. Namun, sulit untuk menangani informasi yang berlebihan dan kumpulan data yang sangat besar (Affonso, 2015).

Berdasarkan uraian di atas, maka penelitian ini menerapkan metode *artificial neural network* dalam memprediksi karakteristik yang mempengaruhi hasil diabetes atau non diabetes, diantaranya yaitu: masa kehamilan, glukosa, tekanan darah, ketebalan kulit, insulin, Indeks massa tubuh (BMI: berat badan dalam kilogram dibagi tinggi badan dalam meter kuadrat), faktor keturunan (diabetes dalam riwayat keluarga), dan umur. Data pasien yang di inputkan akan diproses untuk mendapatkan perhitungan dan mendapatkan hasil prediksi diabetes tersebut.

## **1.2 Pernyataan Masalah**

Berdasarkan latar belakang masalah yang sudah dijelaskan sebelumnya, pernyataan masalah yang diambil dari penelitian ini yaitu :

1. Berapa nilai presisi, akurasi, dan *recall* dari metode *Artificial Neural Network* dalam mendeteksi penyakit diabetes?
2. Aspek atau variabel-variabel apa saja yang mempengaruhi terhadap nilai presisi, akurasi, dan *recall* dalam mendeteksi penyakit diabetes?

### **1.3 Tujuan Penelitian**

Tujuan Penelitian yang diambil dari pernyataan masalah yaitu:

1. Mengukur nilai presisi, akurasi, dan *recall* dari metode *Artificial Neural Network* dalam memprediksi penyakit diabetes.
2. Menganalisis aspek atau variabel-variabel yang mempengaruhi terhadap nilai, presisi, akurasi, dan *recall* dalam memprediksi penyakit diabetes.

### **1.4 Batasan Masalah**

Batasan masalah yang digunakan pada penelitian ini yaitu:

1. Data karakteristik didapatkan dari situs web resmi “*Kaggle*”, seperti masa kehamilan, glukosa, tekanan darah, ketebalan kulit, insulin, Indeks massa tubuh (BMI: berat badan dalam kilogram dibagi tinggi badan dalam meter kuadrat), faktor keturunan seperti diabetes dalam riwayat keluarga, dan umur.
2. Obyek penyakit yang diteliti yaitu penyakit diabetes.

### **1.5 Manfaat Penelitian**

Dari penelitian ini diharapkan bermanfaat bagi:

1. Bagi para tenaga dokter dalam memprediksi penyakit diabetes berdasarkan gejala dari data yang tersedia.

2. Bagi Pusat Kesehatan Masyarakat dalam membantu pasien dalam mendiagnosa penyakit diabetes berdasarkan gejala atau data yang tersedia.
3. Bagi Fakultas Kesehatan dalam menambah pengetahuan dengan adanya sistem prediksi ini sehingga kedepannya bisa dikembangkan.

## BAB II

### STUDI PUSTAKA

#### 2.1 Deteksi Penyakit Diabetes

Deteksi merupakan keingintahuan manusia terhadap hal-hal tertentu, yang mengarah pada pengembangan alat yang membantu memecahkan masalah keingintahuan manusia. Penggunaan atau pengembangan alat deteksi dapat memberikan informasi mengenai penyebab dan gejala kerusakan. Deteksi di bidang teknis membantu menentukan kerusakan alat. Deteksi dapat diartikan sebagai diagnosis dalam bidang kesehatan (Syahrizal & Muhammad, 2018). Deteksi ini melakukan suatu proses untuk memeriksa atau melakukan pemeriksaan terhadap sesuatu dengan menggunakan cara dan teknik tertentu. Deteksi dapat digunakan untuk berbagai masalah, misalnya dalam sistem pendeteksi suatu penyakit, dimana sistem mengidentifikasi masalah-masalah yang berhubungan dengan penyakit yang biasa disebut gejala. Tujuan dari deteksi adalah memecahkan suatu masalah dengan berbagai cara tergantung metode yang diterapkan sehingga menghasilkan sebuah solusi (Wafiah *et al.*, 2013).

Abedini *et al.* (2020) melakukan penelitian menggunakan *Ensemble of Decision Tree, Logistic Regression dan Neural Network* dalam mengklasifikasikan diabetes. Dalam penelitian ini, peneliti mengusulkan metode pembelajaran mesin untuk menganalisis kumpulan data PIMA dan memprediksi risiko diabetes berdasarkan fitur yang tersedia. Peneliti mengusulkan metode klasifikasi ansambel, yang terdiri dari dua level: pada level pertama kami melatih Regresi Logistik dan

Pohon Keputusan (ID3) secara mandiri. Pada level berikutnya, peneliti menggabungkan output dari level sebelumnya dengan menggunakan JST. Dalam percobaan peneliti, menunjukkan bahwa, model ansambel meningkatkan akurasi, dibandingkan dengan metode individual. Hasil percobaan menunjukkan bahwa model yang diusulkan dapat secara efektif menggabungkan model Pohon Keputusan dan Regresi dan meningkatkan akurasi secara keseluruhan. Dibandingkan dengan metode lain dalam literatur, model yang kami usulkan juga menunjukkan peningkatan kinerja.

Penelitian yang dilakukan oleh Rajni dan Amandeep (2019) dalam memprediksi diabetes menggunakan algoritma *RB-bayes* dalam kumpulan data PIMA india. Meskipun ada sejumlah solusi yang tersedia untuk ekstraksi informasi dari kumpulan data yang sangat besar dan untuk memprediksi kemungkinan menderita diabetes, tetapi keakuratan proses penambangannya jauh dari akurat. Untuk mencapai akurasi tertinggi, masalah probabilitas nol yang umumnya dihadapi oleh analisis *naïve bayes* perlu ditangani dengan tepat. Kerangka kerja yang diusulkan *RB-Bayes* bertujuan untuk mengekstraksi informasi yang diperlukan dengan akurasi tinggi yang dapat bertahan dari masalah probabilitas nol dan juga mengonfigurasi akurasi dengan metode lain seperti *Support Vector Machine, Naive Bayes, dan K Nearest Neighbor*. Peneliti menghitung rata-rata untuk menangani data yang hilang dan menghitung probabilitas untuk ya (positif) dan tidak (negatif). Nilai tertinggi antara ya dan tidak menentukan nilai tuple. Ini sebagian besar digunakan dalam klasifikasi teks. Hasil pada dataset diabetes India Pima menunjukkan bahwa metodologi yang diusulkan meningkatkan presisi

sebagai kontras dengan prosedur yang diatur lainnya. Akurasi metodologi dataset besar yang diusulkan adalah 72,9%.

Chang *et al.* (2022) melakukan penelitian algoritma *machine learning* dalam mengklasifikasikan diabetes melitus Pima Indian. Dalam penelitian ini, menggambarkan penggunaan tiga model ML terbimbing yang dapat ditafsirkan: pengklasifikasi *Naïve Bayes*, pengklasifikasi *random forest*, dan J48 decision tree untuk dilatih dan diuji menggunakan dataset diabetes Pima Indians dalam bahasa pemrograman R. Kinerja masing-masing algoritma dianalisis untuk menentukan algoritma dengan akurasi, presisi, sensitivitas, dan spesifisitas terbaik. Penilaian proses keputusan juga dilakukan untuk memperbaiki model. Dapat disimpulkan bahwa model *Naïve Bayes* bekerja dengan baik dengan pemilihan fitur yang lebih baik untuk klasifikasi biner, sedangkan hutan acak bekerja lebih baik dengan lebih banyak fitur. Model *Naïve Bayes* pada subset data 3-faktor menunjukkan hasil yang sama baiknya dengan model hutan acak pada set data lengkap dengan akurasi 79,13% dibandingkan dengan 79,57%, yang merupakan akurasi tertinggi dalam percobaan ini. Kita dapat menyimpulkan bahwa model *Naïve Bayes* bekerja dengan baik dengan pemilihan fitur yang lebih baik untuk klasifikasi biner tetapi gagal dengan banyak fitur yang berkorelasi, sementara hutan acak bekerja lebih baik dengan lebih banyak fitur. Model pohon keputusan J48 secara konsisten dilakukan dengan kisaran tingkat sensitivitas 88,43% (kumpulan data lengkap) hingga 89,92% (subset data 5 faktor), menunjukkan bahwa model ini bagus dalam memprediksi keberadaan diabetes tidak peduli berapa banyak fitur yang harus bekerja. dengan.

Erlin *et al.* (2022) melakukan penelitian mendeteksi penyakit diabetes menggunakan algoritma *Logistic Regression*. Eksperimen dilakukan dengan menggunakan data dari Pima Indians Diabetes Database, yang terdiri dari 768 catatan pasien dengan delapan variabel independen dan satu variabel dependen. Analisis data eksplorasi dilakukan untuk mengekstraksi informasi sebanyak mungkin dari materi khusus menggunakan statistik, dan disajikan menggunakan teknik visual. Beberapa variabel dalam dataset berisi data yang tidak lengkap. Nilai data yang hilang diganti dengan nilai median dari masing-masing variabel. Pengolahan data yang tidak seimbang dilakukan dengan teknik *Synthetic Minority Oversampling Technique* (SMOTE) untuk meningkatkan kelas minoritas menggunakan sampel data sintetik. Model diberi skor berdasarkan matriks kebingungan, yang berkinerja cukup baik dengan skor akurasi 77%. Presisi 75%, penarikan 77%, dan skor F1 76%. Selain itu, makalah ini juga menggunakan teknik *grid search* sebagai pengaturan *hyperparameter* yang dapat meningkatkan kinerja model regresi logistik. Kami menguji dan mengevaluasi kinerja model dasar dengan model setelah menerapkan teknik pencarian *grid*. Hasil eksperimen menunjukkan bahwa model berbasis *hyperparameter tuning* dapat meningkatkan kinerja algoritma regresi logistik prediktif dengan akurasi 82%, akurasi recall 81%, dan skor F1 80%.

Naz dan Ahuja (2020) meneliti terkait prediksi diabetes dengan menggunakan dataset PIMA Indian dalam pendekatan *deep learning*. Dalam penelitian ini, Menghasilkan akurasi yang dicapai oleh pengklasifikasi fungsional *Artificial Neural Network* (ANN), *Naive Bayes* (NB), *Decision Tree* (DT) dan *Deep*

*Learning* (DL) berada dalam kisaran 90–98%. Di antara keempatnya, DL memberikan hasil terbaik untuk onset diabetes dengan tingkat akurasi 98,07% pada dataset PIMA. Oleh karena itu, sistem yang diusulkan ini menyediakan alat prognostik yang efektif bagi petugas kesehatan. Hasil yang diperoleh dapat digunakan untuk mengembangkan alat prognosis otomatis baru yang dapat membantu dalam deteksi dini penyakit ini. Kesimpulan Hasil penelitian menegaskan bahwa DL memberikan hasil terbaik dengan fitur ekstraksi yang paling menjanjikan. DL mencapai akurasi 98,07% yang dapat digunakan untuk pengembangan alat prognosis otomatis lebih lanjut. Keakuratan pendekatan DL selanjutnya dapat ditingkatkan dengan memasukkan data omics untuk prediksi timbulnya penyakit.

Abdi *et al.* (2017) melakukan penelitian mendeteksi penyakit diabetes dengan menggunakan metode *Fuzzy C-Means Clustering* dan *K-Means Clustering*. Dalam penelitian ini menggunakan dataset terbatas pada populasi Pima Indian, hal ini juga dapat menjadi awal yang baik untuk mendeteksi populasi yang lain. Keterbatasan dataset yang diambil pada 1960-1980 menjadi pertimbangan pada hasil clustering, saat ini tes diabetes dilengkapi dengan test urin dan hemoglobin A1c test untuk hasil yang lebih akurat. Penelitian ini melakukan pengelompokan terdiri dari kelompok diabetes dan non-diabetes. Pengujian untuk masing-masing metode dilakukan terhadap 9 data. Hasil pengujian terbaik metode K-Means adalah 73,438% dan untuk metode Fuzzy C-Means adalah 82,812%.

## **2.2 Artificial Neural Network dalam Mendeteksi Penyakit**

Kamlesh *et al.* (2020) melakukan penelitian menggunakan metode Artificial Neural Network dalam memprediksi timbulnya penyakit diabetes. Dalam penelitian ini, sistem diagnosis otomatis diperkenalkan dan dianalisis. Untuk tujuan ini, Buatan Tiga Lapis *Artificial Neural Network* dan dataset Pima Indians Diabetes digunakan. Dalam model prediksi berbasis ANN ini, fungsi aktivasi logistik untuk aktivasi neuron, dan metode Quasi Newton digunakan sebagai algoritma untuk pelatihan. Akibatnya kumulatif gain plot dan sebagai ukuran kualitas model ini skor gain maksimum digunakan. Analisis ini menunjukkan bahwa *Artificial Neural Network* masuk akal untuk meramalkan awal diabetes. Itu kemungkinan pengaturan dalam persiapan dan ujian tes ada di sekitar 55 persen. Penelitian ini juga menunjukkan pendekatan realistis untuk penggunaan jaringan saraf sebagai instrumen pemodelan. Saat ini titik waktu, pengujian terukur tidak cocok. Menuju itu jalan, eksplorasi lebih lanjut diperlukan, dan pekerjaan penelitian ini dapat diperluas untuk membuat dan membandingkan model yang disempurnakan.

Penelitian yang dilakukan oleh Soltani dan Jafarian (2016) dalam mendiagnosis penyakit diabetes menggunakan pendekatan jaringan syaraf tiruan baru menggunakan jaringan saraf tiruan probabilistik untuk pendekatan diagnosis penyakit diabetes tipe II. Penelitian ini memanfaatkan dataset Diabetes Pima Indian dengan 768 sampel dalam percobaan kami. Menurut dataset ini, PNN diimplementasikan di MATLAB. Selain itu, memaksimalkan akurasi diagnosis penyakit Diabetes tipe II dalam pelatihan dan pengujian dataset Diabetes Pima Indian adalah ukuran kinerja dalam makalah ini. Akhirnya, peneliti menyimpulkan

bahwa akurasi pelatihan dan akurasi pengujian dari metode yang diusulkan masing-masing adalah 89,56% dan 81,49%.

Penelitian yang dilakukan Nurhafifah *et al.* (2021) dalam mendeteksi penyakit hipertensi dengan mengimplementasikan metode *Artificial Neural Network* (ANN) yang mana dalam penelitiannya mendeteksi hipertensi berdasarkan ciri-ciri dan atau kebiasaan seseorang. Data hipertensi diperoleh dengan kuesioner kemudian dipraproses dengan melakukan imputasi pada missing value serta transformasi data untuk mengubah data pada range yang sama. Hasil praproses digunakan untuk melatih model ANN dengan 25 input, 3 hidden neuron serta 1 output. Dengan learning rate 0.25 serta epoch sebanyak 1000, hasil percobaan menunjukkan akurasi hingga 96% dengan pembagian data latih sebesar 80% dan data uji sebesar 20%.

Suwarno (2022) melakukan penelitian pengimplementasian Neural Network dalam mendeteksi penyakit diabetes pada renopati. Dalam penelitian ini mendeteksi gejala penyakit ini lebih dini dengan cara menerapkan metode yang dapat menghasilkan akurasi yang tinggi dengan bantuan komputer melalui metode neural network. Data penelitian diperoleh penulis dari situs pengkodean Kaggle (<https://www.kaggle.com>) dimana setiap gambar mempunyai klasifikasi, yaitu: mata normal, mata dengan ringan, mata dengan sedang, mata dengan berat, dan mata dengan sangat berat. Data tersebut selanjutnya diproses dengan menggunakan tools MATLAB untuk preprocessing gambar dan Tensorflow yang digunakan untuk mencari bobot akurasi. Hasil dari penelitian menunjukkan bahwa dengan 1200 kali eksperimen dihasilkan bobot akurasi rata-rata sebesar 88.61% dengan standard

deviasi sebesar 6.41% dengan range bobot akurasi tertinggi sebesar 93.01% dan bobot akurasi terendah sebesar 50.18%.

Irmawati *et al.* (2022) melakukan penelitian pengimplementasian metode ANN untuk mendeteksi penyakit hati, Dalam penelitian ini menghasilkan klasifikasi dengan bantuan algoritma data mining. Dalam makalah ini, Pasien Hati telah diselidiki menggunakan model *Artificial Neural Network* untuk memprediksi seseorang Pasien Hati atau tidak dan analisis menggunakan ANN dengan Python digunakan untuk mengetahui pengaruh variabel input berdasarkan data dalam literature dan mendapat hasil akurasi sebesar 74%.

Niendy dan Ratnadewi (2021) melakukan penelitian penggunaan *Artificial Neural Network* untuk mendeteksi penyakit jantung Aritmia Supraventrikular Dalam penelitian ini dilakukan pemrosesan sinyal elektrokardiogram (EKG) terhadap penyakit Aritmia su-praventrikular perlu dilakukan untuk mendeteksi lebih awal adanya permasalahan pada jantung khususnya penyakit aritmia supraventrikular. *Artificial Neural Network* digunakan untuk mendeteksi penyakit jantung Aritmia supraventrikular dan jantung normal karena kelebihanannya dalam mengklasifikasi suatu data dengan tepat, proses yang singkat dan pengelolaan mandiri. Hasil akhir dalam penelitian ini didapatkan nilai tertinggi dalam keberhasilan mengklasifikasi berasal dari struktur algoritma Multi-Layer Perceptron. Nilai akurasi hasil pengujian tertinggi berasal dari metode pelatihan menggunakan Resilient Backpropagation yaitu sebesar 87,5%. Nilai specificity hasil pengujian tertinggi berasal dari metode pelatihan menggunakan Levenberg

Marquard sebesar 83,3%. Nilai sensitivity hasil pengujian tertinggi berasal dari metode pelatihan menggunakan Resilient Backpropagation yaitu sebesar 100%.

Nurul *et al.* (2018) melakukan penelitian yaitu menerapkan metode *Artificial Neural Network* dalam mendeteksi serangan jantung di rumah sakit awal brok Bekasi. Dalam penelitian ini mendeteksi lebih awal pasien dengan suspect (dugaan) penyakit jantung. Hasil deteksi awal ini dapat digunakan oleh para petugas medis sebagai alat bantu dalam penentuan penyakit jantung dan langkah awal penanganannya. Pendeteksian serangan jantung dapat dilakukan dengan teknik komputasi cerdas menggunakan algoritma Back Propagation Neural Network (BPNN). Hasil penelitian menunjukkan bahwa nilai akurasi hasil prediksi menggunakan algoritma BPNN adalah 96,47%. Sehingga dapat dikatakan bahwa algoritma *Artificial Neural Network* memiliki hasil prediksi yang baik dalam mendeteksi serangan jantung seorang pasien.

Purwono *et al.* (2022) melakukan penelitian pemanfaatan algoritma machine learning *Artificial Neural Network* dalam memprediksi otomatis jenis penyakit hipertensi. Dalam penelitian ini menggunakan teknik optimasi hyperparameter gridsearch cv pada model *Artificial Neural Network*. Parameter yang digunakan dalam model ini yaitu solver='lbfgs', alpha=1e-5, hidden\_layer\_sizes = (5, 2), random\_state = 1. Nilai akurasi ketepatan prediksi dalam menentukan jenis hipertensi ini sebesar 85% yang menunjukkan bahwa model yang dibangun ternyata sudah cukup baik.

Tabel 2.1 merupakan tabel yang berisi perbandingan antara penelitian terdahulu dengan data yang sama dengan penelitian yang dilakukan.

Tabel 2. 1 Perbedaan dengan peneliti terdahulu

No.	Peneliti (Tahun)	Metode	Hasil Penelitian	Perbedaan
1.	Kamlesh <i>et al.</i> (2020)	<i>Artificial Neural Network</i> menggunakan metode <i>Quasi Newton</i>	Dengan menambahkan metode Quasi Newton ini, kumulatif gain plot dan sebagai ukuran kualitas model ini skor gain maksimum digunakan. Analisis ini menunjukkan bahwa ANN ( <i>Artificial Neural Network</i> ) masuk akal untuk meramalkan awal diabetes. Itu kemungkinan pengaturan dalam persiapan dan ujian tes ada di sekitar 55 persen. Penelitian ini juga menunjukkan pendekatan realistis untuk penggunaan jaringan saraf sebagai instrumen pemodelan.	Dalam penelitian ini ditambahkan algoritma metode Quasi Newton dan di tetapkan epoch sebanyak 151
2.	Soltani dan Jafarian (2016)	<i>Artificial Neural Network</i> probabilistik	Peneliti menyimpulkan bahwa akurasi pelatihan dan akurasi pengujian dari metode yang diusulkan masing-masing adalah 89,56% dan 81,49%.	Dalam penelitian ini, menggunakan pendekatan jaringan syaraf tiruan baru menggunakan jaringan saraf tiruan probabilistik untuk pendekatan diagnosis penyakit diabetes tipe II
3.	Abedini <i>et al.</i> (2020)	<i>Ensemble of Decision Tree, Logistic Regression dan Neural Network</i>	Hasil percobaan menunjukkan bahwa model yang diusulkan dapat secara efektif menggabungkan model Pohon Keputusan dan Regresi dan meningkatkan akurasi secara keseluruhan.	Metode yang digunakan pada penelitian ini yaitu <i>Ensemble of Decision Tree, Logistic Regression dan Neural Network</i>
4.	Rajni dan Amandeep (2019)	algoritma <i>RB-bayes</i>	Hasil pada dataset diabetes India Pima menunjukkan bahwa metodologi yang diusulkan meningkatkan presisi sebagai kontras dengan prosedur yang diatur lainnya. Akurasi metodologi dataset besar yang diusulkan adalah 72,9%.	Penelitian ini menggunakan metode <i>RB-Bayes</i>

Lanjutan tabel Perbedaan Dengan Penelitian Terdahulu

No.	Peneliti (Tahun)	Metode	Hasil Penelitian	Perbedaan
5.	Chang <i>et al.</i> (2022)	Dalam penelitian ini, menggambarkan penggunaan tiga model ML terbimbing yang dapat ditafsirkan: pengklasifikasi <i>Naïve Bayes</i> , pengklasifikasi <i>random forest</i> , dan <i>J48 decision tree</i>	Model <i>Naïve Bayes</i> pada subset data 3-faktor menunjukkan hasil yang sama baiknya dengan model hutan acak pada set data lengkap dengan akurasi 79,13% dibandingkan dengan 79,57%, yang merupakan akurasi tertinggi dalam percobaan ini. Model pohon keputusan <i>J48</i> secara konsisten dilakukan dengan kisaran tingkat sensitivitas 88,43% (kumpulan data lengkap) hingga 89,92% (subset data 5 faktor).	Dalam penelitian ini menggunakan pemrograman Bahasa R dan menggunakan tiga model <i>machine learning</i> yaitu <i>Naïve Bayes</i> , <i>random forest</i> , dan <i>J48 decision tree</i> .
6.	Naz dan Ahuja (2020)	<i>Artificial Neural Network</i> , <i>Naive Bayes</i> , <i>Decision Tree</i> dan <i>Deep Learning</i>	Di antara keempatnya, DL memberikan hasil terbaik untuk onset diabetes dengan tingkat akurasi 98,07% pada dataset PIMA. Oleh karena itu, sistem yang diusulkan ini menyediakan alat prognostik yang efektif bagi petugas kesehatan..	Dalam penelitian ini menggunakan metode <i>Artificial Neural Network</i> , <i>Naive Bayes</i> , <i>Decision Tree</i> dan <i>Deep Learning</i>
7.	Abdi <i>et al.</i> (2017)	<i>Fuzzy C-Means Clustering</i> dan <i>K-Means Clustering</i>	Penelitian ini melakukan pengelompokan terdiri dari kelompok diabetes dan non-diabetes. Pengujian untuk masing-masing metode dilakukan terhadap 9 data. Hasil pengujian terbaik metode <i>K-Means</i> adalah 73,438% dan untuk metode <i>Fuzzy C-Means</i> adalah 82,812%.	Dalam penelitian ini menggunakan metode <i>Fuzzy C-Means Clustering</i> dan <i>K-Means Clustering</i>

## BAB III

### METODOLOGI PENELITIAN

#### 3.1 Pengumpulan Data

Data yang digunakan pada penelitian ini yaitu data yang berasal dari Institut Nasional Diabetes dan Pencernaan dan Penyakit Ginjal yang diperoleh dari situs resmi *Kaggle* "<https://www.kaggle.com/>" yang dapat diakses oleh siapapun. Dalam data tersebut berjumlah 768 data pasien berusia minimal 21 tahun. Data tersebut berupa masa kehamilan, glukosa yang menyatakan tingkat glukosa dalam darah, tekanan darah, ketebalan kulit, insulin dalam darah, Indeks massa tubuh (BMI: berat badan dalam kilogram dibagi tinggi badan dalam meter kuadrat), faktor keturunan (diabetes dalam riwayat keluarga), serta umur. Pada data ini terbagi menjadi data latih dan data uji. Untuk jumlah keseluruhan data ini terdiri dari 768 baris *record* dan 8 fitur parameter. Fitur dataset yang digunakan terdapat pada tabel 3.1.

Tabel 3.1 Fitur Dataset

No	Parameter	Keterangan
1	<i>Pregnancies</i>	Kolom yang menyatakan masa kehamilan
2	<i>Glucose</i>	Kolom yang menyatakan kadar Glukosa dalam darah
3	<i>BloodPressure</i>	Kolom yang menyatakan pengukuran tekanan darah
4	<i>SkinThickness</i>	Kolom yang menyatakan ketebalan kulit
5	<i>Insulin</i>	Kolom yang menyatakan kadar Insulin dalam darah
6	<i>BMI</i>	Kolom yang menyatakan indeks massa tubuh
7	<i>DiabetesPedigreeFunction</i>	Kolom yang menyatakan persentase Diabetes
8	<i>Age</i>	Kolom yang menyatakan usia

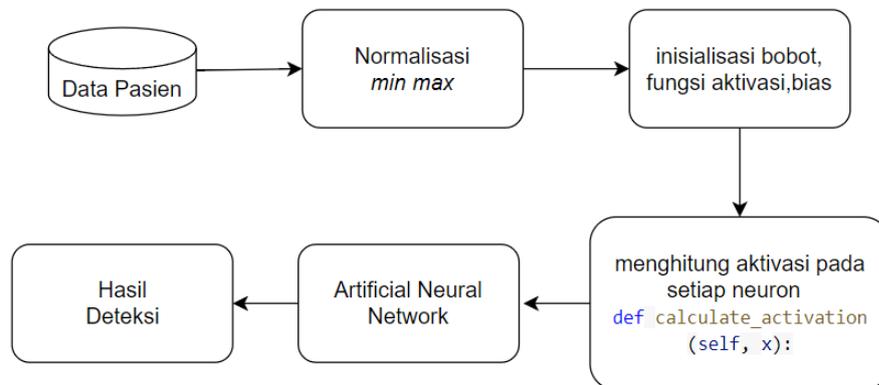
Pada Tabel 3.2 yaitu sebagai contoh gambaran dataset yang akan digunakan sebagai input, dataset penyakit diabetes memiliki 8 kolom yang berupa faktor risiko, seperti yang telah dijelaskan pada Tabel 3.1.

Tabel 3.2 Contoh Dataset Penyakit Diabetes

<i>Pregnancies</i>	<i>Glucose</i>	<i>Blood Pressure</i>	<i>Skin thickness</i>	<i>Insulin</i>	<i>BMI</i>	<i>DPF</i>	<i>Age</i>
6	148	72	35	0	33,6	0,627	50
1	85	66	29	0	26,6	0,351	31
8	183	64	0	0	23,3	0,672	32
1	89	66	23	94	28,1	0,167	21
0	137	40	35	168	43,1	2,288	33
5	116	74	0	0	25,6	0,201	30
3	78	50	32	88	31	0,248	26
10	115	0	0	0	35,3	0,134	29
2	197	70	45	543	30,5	0,158	53
8	125	96	0	0	0	0,232	54
4	110	92	0	0	37,6	0,191	30
10	168	74	0	0	38	0,537	34
10	139	80	0	0	27,1	1,441	57

### 3.2 Desain Sistem

Dalam penelitian ini dibuat sebuah alur desain sistem prediksi penyakit diabetes yang bisa dilihat pada Gambar 3.1. Terdapat beberapa proses dalam alur sistem ini yaitu preprocessing data, pelatihan data, percobaan model, dan perhitungan error.



Gambar 3.1 Desain Sistem

Pada tahap *preprocessing* ini, data pasien yang sudah didapatkan akan melewati proses normalisasi data *min max*, kemudian menginisialisasi bobot, fungsi aktivasi, dan bias, kemudian membangun model Artificial Neural Network dan menghasilkan hasil deteksi.

Pada Gambar 3.1 diatas dapat dijelaskan bahwa normalisasi data *min-max* merupakan metode normalisasi dengan melakukan transformasi linier terhadap data asli sehingga menghasilkan keseimbangan nilai perbandingan antar data saat sebelum dan sesudah proses. Dalam penggunaan normalisasi ini digunakan rumus sebagai berikut (Arhami dan Nasir, 2020):

$$X = \frac{x_n - \text{Min}}{\text{Max} - \text{Min}} \quad (3.1)$$

Keterangan:

- $x_n$  = nilai dari dataset yang ingin dinormalisasi.
- min = nilai terkecil dalam dataset  $x$ .
- max = nilai terbesar dalam dataset  $x$ .

Tabel 3.3 Contoh data pasien sebelum dilakukan normalisasi

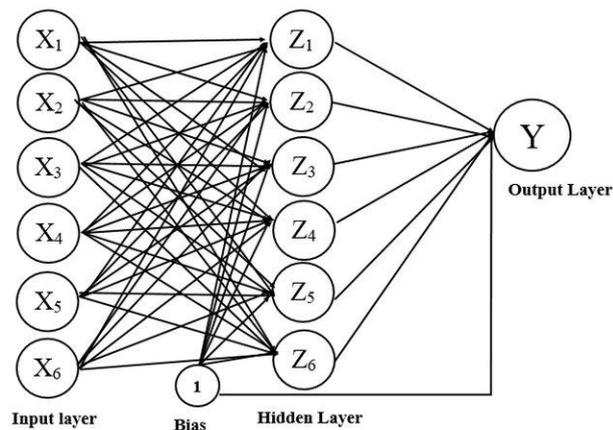
<i>Pregnancies</i>	<i>Glucose</i>	<i>Blood Pressure</i>	<i>Skin thickness</i>	<i>Insulin</i>	<i>BMI</i>	<i>Diabetes Pedigree Function</i>	<i>Age</i>
6	148	72	35	0	33.6	0.627	50
1	85	66	29	0	26.6	0.351	31
8	183	64	0	0	23.3	0.672	32
1	89	66	23	94	28.1	0.167	21
0	137	40	35	168	43.1	2.288	33
5	116	74	0	0	256	0.201	30
3	78	50	32	88	31	0.248	26
10	115	0	0	0	35.3	0.134	29
2	197	70	45	543	30.5	0.158	53
8	125	96	0	0	0	0.232	54

Pada Tabel 3.3 dapat dilihat nilai asli dari dataset pasien sebelum dilakukannya preprocessing. Banyak data yang memiliki rentang berbeda sehingga harus dilakukan normalisasi. Dataset tersebut ditransformasi menggunakan metode min-max normalization dengan mengolah nilai minimum dan maksimum dari setiap atribut. Rentang yang digunakan dalam metode ini adalah 0-1. Rumus *min-max normalization* dapat digunakan untuk melakukan normalisasi dengan metode min-max normalization. Nilai yang dihasilkan setelah pengolahan memiliki rentang nilai yang seimbang.

### 3.3 Artificial Neural Network

Suatu pendekatan model kecerdasan yang diilhami dari struktur otak manusia dan kemudian diimplementasikan menggunakan program komputer yang mampu menyelesaikan sejumlah proses perhitungan selama proses learning berlangsung. Berdasarkan penjelasan tersebut ANN dapat direpresentasikan

menjadi 3 bagian yang terdiri dari layer masukan (*input*) dan layer keluar (*output*), dan layer tersembunyi yang memproses masukan dari layer masukan menjadi sesuatu yang dapat diterima layer keluaran.



Gambar 3.2 Arsitektur Artificial Neural Network (Pohan *et al.* 2020)

Keterangan:

$X$  = neuron pada *input* layer

$Z$  = neuron pada *hidden* layer

$Y$  = *output* layer

1 = *bias*

$w$  = bobot yang berada di antara *hidden* dan *output* layer

Dalam penelitian ini, dapat dilihat pada Gambar 3.4 memiliki dua layer yaitu input layer, hidden layer, dan output layer. Input layer tersebut terdiri dari node-node yang akan diinputkan data-data yang sudah di preprocess sebelumnya. Backpropagation pada jaringan saraf tiruan menjelaskan bagaimana lapisan input menerima data dari vektor input dan mengirimkannya ke lapisan tersembunyi ( $Z_i$ ,  $Z_j$ , dan  $Z_q$ ). Kemudian, unit-unit di lapisan tersembunyi ( $Z_i$ ,  $Z_j$ , dan  $Z_q$ ) menerima bobot input dan mentransfer bobot tersebut ke lapisan output ( $Y$ ) dengan fungsi transfer. Informasi tersebar ke dalam jaringan dengan setiap manipulasi input, dan hasil dihitung di setiap unit proses. Setiap unit di lapisan output menghitung bobot

dan kesalahan. Kemudian, penyebaran balik dari lapisan output kembali ke lapisan tersembunyi untuk melihat kesalahan dan nilai bobot yang dihitung, lalu kembali ke lapisan tersembunyi untuk memperbarui bobot.

Setiap neuron yang terdapat pada input layer memiliki bobot yang disimbolkan dengan  $w$ , kedua nilai ini akan dioperasikan menggunakan weighted sum model dimana bobot akan dikalikan dengan input. Total penjumlahan antara bobot dikalikan dengan input, selanjutnya akan dijumlahkan dengan nilai bias. Hasil operasi ini selanjutnya dimasukkan ke dalam fungsi aktivasi sigmoid, hasil dari proses ini disimbolkan dengan  $h$ .

- a. Input layer, pada lapisan ini berisikan neuron - neuron yang digunakan untuk menerima input untuk proses learning dan recognition dari ANN. Dalam neuron tersebut berisikan data yang telah didapat sebelumnya, yaitu masa kehamilan, glukosa yang menyatakan tingkat glukosa dalam darah, tekanan darah, ketebalan kulit, insulin dalam darah, Indeks massa tubuh (BMI: berat badan dalam kilogram dibagi tinggi badan dalam meter kuadrat), faktor keturunan (diabetes dalam riwayat keluarga), serta umur.
- b. Hidden layer, pada lapisan ini berisikan neuron - neuron yang digunakan untuk mengganti kontribusi yang digunakan oleh neuron pada output layer.
- c. Output Layer, pada lapisan ini berisikan neuron - neuron yang akan bereaksi terhadap data input yang diberikan.

Dalam jaringan feedforward, sinyal input dipropagasikan dari lapisan input melalui lapisan tersembunyi dan kemudian ke lapisan output tanpa adanya koneksi mundur atau umpan balik. Dalam jaringan feedforward, setiap neuron pada

lapisan tertentu terhubung dengan setiap neuron pada lapisan berikutnya, tetapi tidak ada koneksi mundur ke lapisan sebelumnya. Informasi mengalir hanya dari input ke output melalui lapisan-lapisan tersembunyi.

Setiap neuron pada lapisan tertentu menerima sinyal masukan dari neuron-neuron pada lapisan sebelumnya dan menghasilkan sinyal keluaran yang kemudian diteruskan ke semua neuron pada lapisan berikutnya. Setiap koneksi antara neuron memiliki bobot yang mempengaruhi seberapa besar kontribusi setiap neuron terhadap keluaran. Tujuan dari jaringan feedforward adalah untuk mempelajari fungsi yang dapat memetakan input ke output yang diinginkan. Ini dapat dilakukan dengan mengatur bobot koneksi antara neuron selama proses pelatihan jaringan, menggunakan algoritma seperti backpropagation.

Dalam penelitian ini fungsi feed forward digunakan untuk menghitung output dari setiap layer pada model. Fungsi ini mengambil input data  $x$  sebagai parameter dan melewatkannya melalui setiap layer menggunakan fungsi aktivasi pada masing-masing layer. Hasil output dari layer saat ini menjadi input pada layer berikutnya. Pada setiap iterasi, fungsi ini juga mencetak nilai  $x$  setelah melewati layer tersebut. Hal ini dapat membantu kita memahami bagaimana input diubah menjadi output pada setiap layer. Untuk menghitung hidden layer diperlukan persamaan fungsi aktivasi 3.2.

$$h_j = (\sum x_i * w_{ij}) + b \quad (3.2)$$

Keterangan:

- $h_j$  = output atau aktivasi dari neuron ke-j.
- $x_i$  = nilai input dari neuron ke-i pada lapisan sebelumnya.
- $w_{ij}$  = bobot sinapsis (weight) yang menghubungkan neuron ke-i.
- $b$  = bias yang ditambahkan ke total.

Fungsi aktivasi merupakan fungsi yang digunakan pada jaringan saraf untuk mengaktifkan atau tidak mengaktifkan neuron. Fungsi aktivasi Sigmoid membuat output dari lapisan tersebut selalu bernilai antara 0 dan 1, yang membuatnya ideal untuk digunakan sebagai lapisan keluaran pada model klasifikasi biner, seperti memprediksi apakah pasien memiliki diabetes atau tidak. Fungsi aktivasi dapat berupa fungsi linear, threshold, atau sigmoid. Fungsi sigmoid biasa digunakan karena fungsi ini mengombinasikan nearly linear behavior, curvilinear behavior, dan nearly constant behavior. Fungsi aktivasi sigmoid dapat diketahui sebagai berikut:

$$\text{sigmoid}(x) = \frac{1}{(1+e^{-x})} \quad (3.3)$$

Keterangan:

$\text{sigmoid}(x)$  = nilai output dari fungsi sigmoid untuk input  $x$ .

$e$  = bilangan Euler (2.71828).

$e^{-x}$  = eksponensial dari  $-x$ .

Algoritma yang digunakan untuk melatih jaringan saraf tiruan (ANN) dengan menghitung gradient loss function terhadap parameter-parameter model, seperti bobot (weights) dan bias (biases). Dengan melakukan ini, kita dapat melakukan penyesuaian bobot dan bias dalam model sehingga dapat lebih cocok dengan data.

Struktur backpropagation merupakan salah satu konsep desain dalam jaringan saraf tiruan yang dapat digunakan untuk memahami dan menganalisis pola data sebelumnya dengan lebih akurat, sehingga menghasilkan keluaran yang lebih tepat dengan tingkat kesalahan minimal.

Proses Training merupakan tahapan yang krusial dalam mengembangkan kemampuan model *Artificial Neural Network* untuk memahami pola dalam data. Pada awalnya, bobot dan bias dalam jaringan diinisialisasi secara acak. Selanjutnya, data pelatihan, yang terdiri dari pasangan input dan output yang diharapkan, diteruskan melalui jaringan. Output yang dihasilkan dibandingkan dengan output yang diharapkan, dan kesalahan prediksi diukur menggunakan fungsi kesalahan. Selama iterasi berulang, jaringan akan memperbarui parameter-nya, memperbaiki kemampuannya dalam menangkap pola-pola yang kompleks. Selain itu, pemilihan fungsi aktivasi dan arsitektur jaringan menjadi faktor penting dalam menentukan kemampuan ANN untuk memodelkan hubungan dalam data. Hasil akhirnya adalah sebuah model ANN yang dapat digunakan untuk melakukan prediksi atau pengklasifikasian pada data baru.

Langkah-langkah dalam membangun algoritma *Artificial Neural Network backpropagation* yakni sebagai berikut (Arhami dan Nasir, 2020):

1. Langkah pertama dalam memulai algoritma backpropagation adalah melakukan inisialisasi bobot, yang melibatkan pengambilan nilai bobot secara acak dengan ukuran yang cukup kecil.
2. Tahap perambatan maju (*forward propagation*)
  - a) Setiap unit input ( $X_i$ ,  $i=1,2,3,\dots,n$ ) menerima sinyal  $x_i$  dan meneruskan sinyal tersebut ke semua unit pada lapisan tersembunyi.
  - b) Setiap unit tersembunyi ( $Z_j$ ,  $j=1,2,3,\dots,p$ ) menjumlahkan bobot sinyal input, ditunjukkan dengan persamaan (3.4).

$$Z_{inj} = v_{j0} + \sum_{i=1}^n v_{ij}x_i \quad (3.4)$$

Keterangan :

- $Z_{inj}$  = Nilai input pada neuron ke- $j$ .  
 $v_{j0}$  = Bias atau ambang batas untuk neuron ke- $j$ .  
 $v_{ij}$  = Bobot sinapsis yang menghubungkan input ke- $i$  dengan neuron ke- $j$ .  
 $x_i$  = Nilai input dari fitur ke- $i$ .  
 $k$  = Jumlah fitur pada input.

Dan menerapkan fungsi aktivasi untuk menghitung sinyal output-nya, ditunjukkan dengan persamaan (3.4).

$$z_j = f(z_{inj}) = \frac{1}{1+e^{-(z_{inj})}} \quad (3.5)$$

Keterangan:

- $z_j$  = aktivasi pada neuron ke- $j$  pada lapisan tersebut.  
 $Z_{inj}$  = nilai total input ke neuron ke- $j$  sebelum fungsi aktivasi diterapkan.  
 $e^{-Z_{inj}}$  = nilai eksponensial dari negatif  $Z_{inj}$ .  
 $f(Z_{inj})$  = fungsi sigmoid yang diterapkan pada  $Z_{inj}$ .

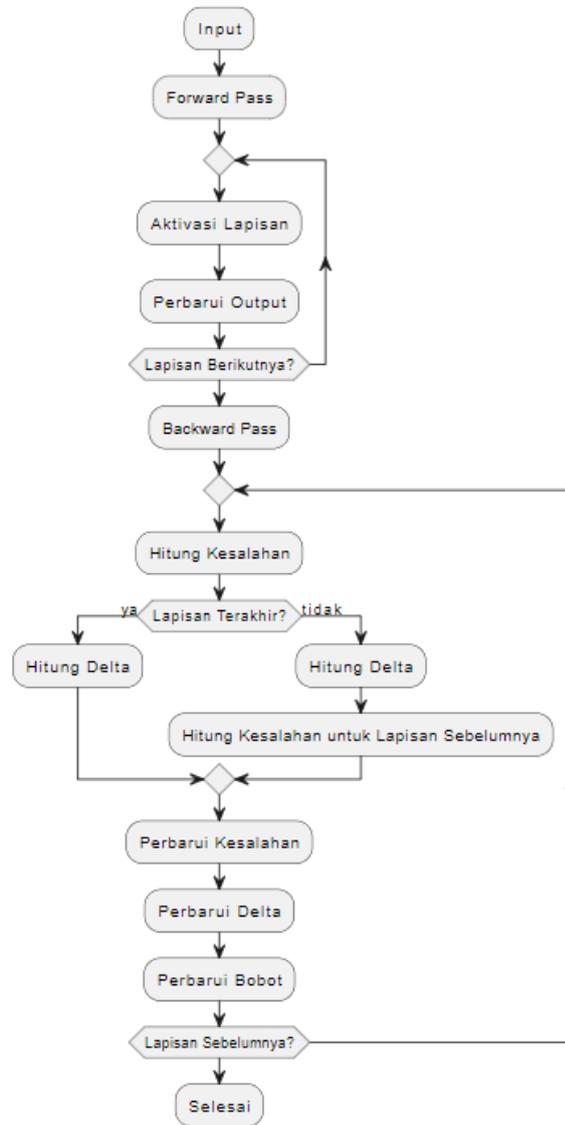
Kemudian Setiap unit output ( $Y_k$ ,  $k=1,2,3,\dots,m$ ) menjumlahkan bobot sinyal input, ditunjukkan dengan persamaan (3.4). Dan menerapkan fungsi aktivasi untuk menghitung sinyal output-nya, ditunjukkan dengan persamaan (3.6).

$$Y_{ink} = w_{0k} + \sum_{j=1}^m w_{jk}z_j \quad (3.6)$$

Keterangan:

- $Y_{ink}$  : Output dari neuron ke- $k$  pada lapisan tersebut.  
 $w_{0k}$  : Bobot asosiasi yang terkait dengan bias dari neuron ke- $k$ .  
 $w_{jk}$  : Bobot yang menghubungkan neuron ke- $j$  pada lapisan sebelumnya dengan neuron ke- $k$ .  
 $z_j$  : Output atau aktivasi dari neuron ke- $j$  pada lapisan sebelumnya.  
 $m$  : Jumlah neuron pada lapisan sebelumnya.

### 3. Tahap *Backpropagation*



Gambar 3.3 *Flowchart Backpropagation*

Setiap unit output ( $Y_k, k=1,2,3,\dots, m$ ) menerima pola target yang sesuai dengan pola input pelatihan, kemudian hitung error, ditunjukkan dengan persamaan (3.7).

$$\delta_k = t_k - y_k f'(y_{in_k}) \quad (3.7)$$

Keterangan:

$\delta_k$  : Kesalahan pada neuron ke- $k$  pada lapisan output.  
 $t_k$  : Nilai target yang diinginkan untuk neuron ke- $k$ .  
 $y_k$  : Output atau aktivasi dari neuron ke- $k$  pada lapisan output.  
 $f'(y_{in_k})$ : Turunan dari fungsi aktivasi ( $f$ ) yang diaplikasikan pada nilai input ( $y_{in_k}$ ) dari neuron ke- $k$ .

$f'$  adalah turunan dari fungsi aktivasi. Kemudian menghitung korelasi bobot, ditunjukkan dengan persamaan (3.8).

$$\Delta w_{jk} = \alpha \delta_k z_j \quad (3.8)$$

Keterangan:

$\Delta w_{jk}$  : Perubahan bobot sinapsis antara neuron ke- $j$  dan neuron ke- $k$ .  
 $\alpha$  : Laju pembelajaran (*learning rate*).  
 $\delta_k$  : Kesalahan pada neuron ke- $k$  pada lapisan output.  
 $z_j$  : Output atau aktivasi dari neuron ke- $j$  pada lapisan sebelumnya.

Dan menghitung koreksi bias, ditunjukkan dengan persamaan (3.9).

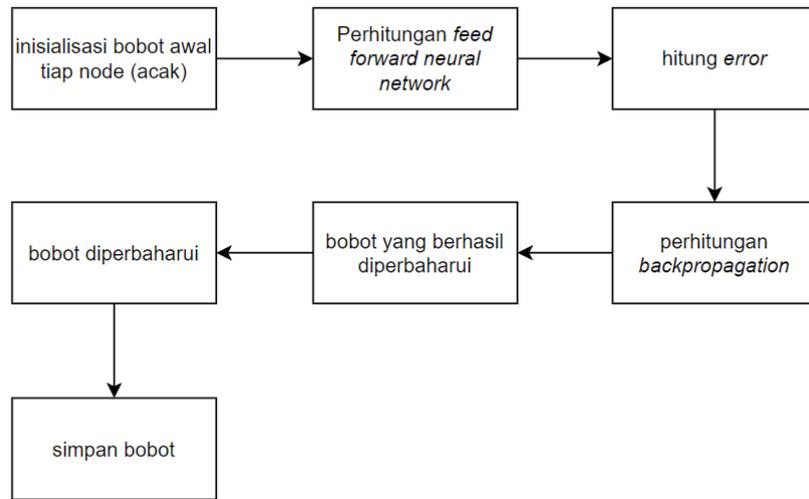
Sekaligus mengirimkan  $\delta_k$  ke unit-unit yang ada di lapisan paling kanan.

$$\Delta w_{0k} = \alpha \delta_k \quad (3.9)$$

Keterangan:

$\Delta w_{0k}$  : Perubahan bobot bias ( $w_{0k}$ ) yang terkait dengan neuron ke- $k$  pada lapisan output.  
 $\alpha$  : Laju pembelajaran (*learning rate*).  
 $\delta_k$  : Kesalahan pada neuron ke- $k$  pada lapisan output.

Selanjutnya, langkah awal adalah menentukan nilai awal untuk variabel-variabel yang penting dalam perhitungan, seperti laju pembelajaran (*learning rate*), dan jumlah maksimum epoch. Selain itu, bobot awal dan bias awal diinisialisasi dengan angka-angka acak yang memiliki magnitudo kecil. Ilustrasi dari proses pelatihan dalam metode *Artificial Neural Network* dapat dilihat pada Gambar 3.3.



Gambar 3. 4 Alur Proses *Training*

Setelah error dihitung, bobot diubah dengan cara mengalikan delta error (perbedaan antara output yang seharusnya dengan output aktual dikalikan dengan turunan fungsi aktivasi) dengan input pada layer sebelumnya. Perubahan bobot tersebut dikalikan dengan learning rate untuk membatasi ukuran langkah yang diambil dalam setiap iterasi dan menghindari overshooting. Pada penelitian ini dilakukan perhitungan error bobot diubah untuk meminimalkan error tersebut. Hal ini dilakukan dengan iterasi melalui layer-layer dalam jaringan dan memperbarui bobot pada setiap layer

Kemudian, pada “Y” sebagai output ini menggunakan rumus sebagai berikut:

$$Y = \max (H_j) \quad (3.10)$$

Keterangan:

- $Y$  = variabel atau simbol yang mewakili nilai maksimum.
- $\text{Max}$  = fungsi yang memberikan nilai maksimum dalam himpunan nilai.
- $H_j$  = sekumpulan nilai, diindeks oleh  $j$ .

Dimana “ $H_j$ ” tersebut adalah rangkaian rumus dari :

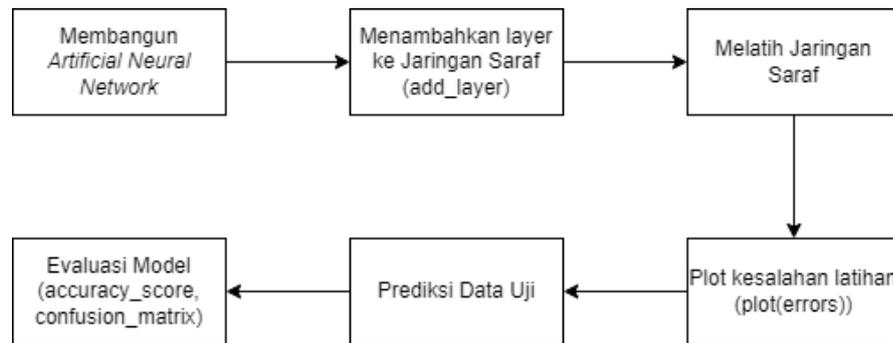
$$H_j = \frac{1}{1+e^{-(\sum x_i * w_{ij}) + b}} \quad (3.11)$$

Keterangan:

- $H_j$  = output dari unit ke-j dalam jaringan saraf.
- $e$  = bilangan Euler, yang merupakan konstanta matematis (sekitar 2.71828).
- $x_i$  = input ke-i untuk unit tersebut.
- $w_{ij}$  = bobot (*weight*) antara input ke-i dan unit ke-j.
- $b$  = bias yang ditambahkan ke hasil penjumlahan.

Setelah melalui tahap pelatihan, bobot-bobot akhir berhasil diperoleh. Bobot-bobot ini kemudian menjadi masukan utama dalam fase pengujian. Tujuan dari proses pengujian algoritma jaringan saraf adalah untuk mengevaluasi kemampuan model yang telah dilatih dalam mengaplikasikan pengetahuannya pada data yang belum pernah dilihat sebelumnya.

Setelah melalui tahap pelatihan, bobot-bobot akhir berhasil diperoleh. Bobot-bobot ini kemudian menjadi masukan utama dalam fase pengujian. Tujuan dari proses pengujian algoritma jaringan saraf adalah untuk mengevaluasi kemampuan model yang telah dilatih dalam mengaplikasikan pengetahuannya pada data yang belum pernah dilihat sebelumnya. Ilustrasi dari alur proses pengujian algoritma jaringan saraf dapat ditemukan pada Gambar 3.5.



Gambar 3. 5 Alur Proses *Testing*

## BAB IV

### UJI COBA DAN PEMBAHASAN

#### 4.1 Skenario Uji Coba

Dalam bab ini akan dibahas untuk mengetahui evaluasi hasil dari deteksi penyakit diabetes menggunakan metode *Artificial Neural Network*. Untuk mendapatkan hasil deteksi diabetes menggunakan metode ANN ini diinputkan data yang telah didapatkan penulis. Hasil deteksi kemudian dilakukan perhitungan menggunakan *confusion matrix* untuk mendapatkan hasil akurasi, presisi, *recall*. Langkah sebelumnya yang digunakan yaitu mulai dari pengumpulan data, *preprocessing*, pembagian dataset, dan proses deteksi menggunakan ANN.

##### 4.1.1 Data Penelitian

Dataset yang digunakan pada penelitian ini dibagi menjadi dua bagian, yaitu data latih dan data uji. Pada data latih tersebut sebanyak 80% dari 768 data, dan data uji 20% dari 768 data. Dataset ini diperoleh dari situs "<https://www.kaggle.com/datasets/mathchi/diabetes-data-set>" yang meliputi masa kehamilan, glukosa yang menyatakan tingkat glukosa dalam darah, tekanan darah, ketebalan kulit, insulin dalam darah, Indeks massa tubuh (BMI: berat badan dalam kilogram dibagi tinggi badan dalam meter kuadrat), faktor keturunan (diabetes dalam riwayat keluarga), serta umur.

Setelah tahap pengumpulan data, langkah selanjutnya yaitu melakukan *preprocessing*, yang mana di tahap ini meliputi normalisasi data, *split* data yang

akan membagi data menjadi data uji dan data latih. Berikut adalah source code dari proses normalisasi *min max*.

```
def minmax_normalize(data):
    min_values = []
    max_values = []
    for i in range(len(data[0])):
        min_values.append(min([row[i] for row in data]))
        max_values.append(max([row[i] for row in data]))
    for i in range(len(data)):
        for j in range(len(data[0])):
            data[i][j] = (data[i][j] - min_values[j]) /
(max_values[j] - min_values[j])
    return data
```

Gambar 4.1 Source Code Normalisasi

Setelah dilakukan normalisasi, maka menghasilkan perubahan nilai setiap fitur seperti pada gambar berikut.

Tabel 4.1 Hasil Normalisasi

<i>Pregnancies</i>	<i>Glucose</i>	<i>Blood Pressure</i>	<i>Skin thickness</i>	<i>Insulin</i>	<i>BMI</i>	<i>Diabetes Pedigree Function</i>
0,352941	0,743719	0,590164	0,353535	0	0,500745	0,234415
0,058824	0,427136	0,540984	0,292929	0	0,396423	0,116567
0,470588	0,919598	0,52459	0	0	0,347243	0,253629
0,058824	0,447236	0,540984	0,232323	0,111111	0,418778	0,038002
0	0,688442	0,327869	0,353535	0,198582	0,642325	0,943638
0,294118	0,582915	0,606557	0	0	0,38152	0,052519
0,176471	0,39196	0,409836	0,323232	0,104019	0,461997	0,072588
0,588235	0,577889	0	0	0	0,52608	0,023911
0,117647	0,98995	0,57377	0,454545	0,641844	0,454545	0,034159
0,352941	0,743719	0,590164	0,353535	0	0,500745	0,234415

Terlihat perubahan nilai pada Tabel 4.1 diatas dengan melakukan penerapan normalisasi dengan nilai yang berkisar antara 0 sampai 1. Setelah dilakukannya proses data ini kemudian melakukan pembagian data menjadi data latih dan data uji (*split data*). Data latih ini nanti nantinya digunakan sebagai pendukung untuk membangun model pelatihan sementara pada bagian data uji akan digunakan sebagai pendukung untuk membangun model pengujian sistem. Berikut *source code* untuk pembagian data latih dan data uji.

```
from sklearn.model_selection import train_test_split
x_train_minmax, x_test_minmax, y_train, y_test =
train_test_split(x_minmax, y, test_size=0.20, random_state=0)
```

Gambar 4.2 Source Code Split Data

Dari data yang ada tersebut, peneliti melakukan pembagian data menjadi data latih dan data uji. Data latih tersebut 80% dari 768 data, dan data uji 20% dari 768 data. Kemudian setelah melakukan split data, yaitu menentukan nilai bias dengan menggunakan rumus seperti pada rumus 3.2 diatas.

#### 4.1.2 Menghitung Kinerja Sistem

Skenario uji sangat penting untuk memastikan bahwa model atau algoritma yang telah dibuat dapat bekerja dengan baik dalam berbagai situasi dan memenuhi tujuan yang diharapkan. Data yang digunakan pada penelitian ini berjumlah 768 data pasien diabetes. Pembagian data dibagi menjadi data testing dan data uji. Data training diharuskan memiliki perbandingan yang lebih besar daripada data testing. Data training berfungsi untuk membuat sebuah mode. Data testing merupakan data yang digunakan untuk menguji model yang dibentuk oleh data training.

Dalam percobaan menggunakan model *Artificial Neural Network*, nilai learning rate harus ditentukan terlebih dahulu untuk mengoptimalkan klasifikasi. Selain itu, nilai epoch juga harus dipilih dengan cermat karena akan mempengaruhi hasil pelatihan data. Pada skenario uji ini, digunakan nilai learning rate terbaik yaitu 0.04. Setelah sistem dibangun, dilakukan evaluasi dengan menghitung akurasi, presisi, dan recall menggunakan persamaan yang telah ditentukan.

Tabel 4.2 Confusion Matrix

Kelas	Diklasifikasikan sebagai	
	Positif	Negatif
+	True Positive (TP)	False Positive (FP)
-	True Negative (TN)	False Negative (FN)

Tabel 4.3 adalah klasifikasi dari confusion matrix. True positive (TP) berarti model klasifikasi dengan benar memberi label jumlah tupel positif. True Negative (TN) berarti bahwa model klasifikasi dengan benar memberi label jumlah tupel negatif. False positive (FP) berarti bahwa model klasifikasi memberi label yang salah untuk jumlah tupel negatif. False Negative (FN) menunjukkan bahwa model klasifikasi memberi label yang salah untuk jumlah tupel positif. Akurasi adalah ukuran kinerja model klasifikasi dan merupakan persentase jumlah data yang diprediksi dengan benar dari total data.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (4.1)$$

Pengukuran performance metrics terdiri dari presisi, dan *recall*. Presisi adalah rasio positif atau derajat keandalan, yaitu proporsi prediksi berlabel positif

yang benar terhadap prediksi positif keseluruhan. Rumus untuk menghitung presisi dapat dilihat pada Persamaan 4.2.

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (4.2)$$

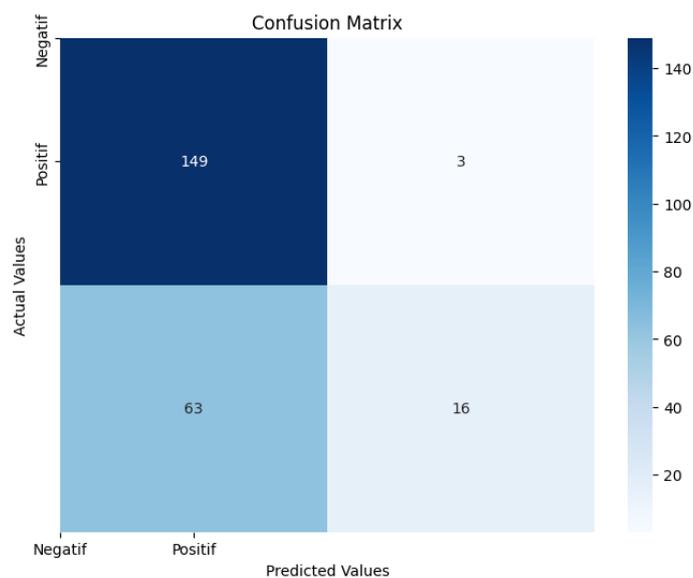
Recall juga dikenal sebagai true positive rate atau sensitivitas. Recall juga disebut sebagai derajat keandalan model dalam mendeteksi data berlabel positif dengan benar. Rumus untuk menghitung recall dapat dilihat pada persamaan 4..

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (4.3)$$

## 4.2 Hasil Uji Coba

### 4.2.1 Uji Coba Ratio Data 70:30

Pengujian pada skenario 1 ini yang mana menggunakan ratio 70:30 yang dilakukan dengan menerapkan *learning rate* sebesar 0.04 dan nilai epoch sebanyak 50. Sehingga diperoleh hasil dari pengujian tersebut menggunakan confusion matrix seperti pada gambar 4.3.



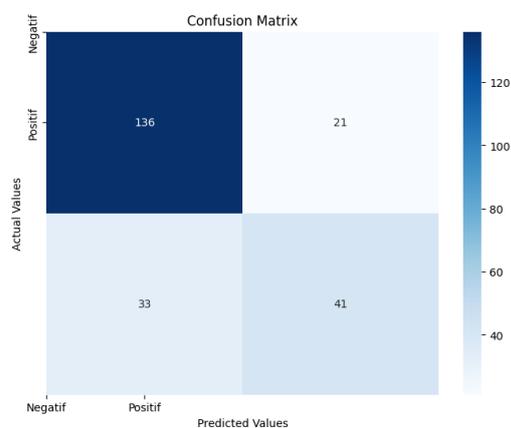
Gambar 4.3 Confusion Matrix 70:30

Pada gambar 4.3 ini menunjukkan hasil dari pengujian skenario 1 yang menggunakan ratio 70:30, *learning rate* 0.04, dengan *epoch* 50. Hasil yang diberikan berhasil diprediksi dengan baik oleh *Artificial Neural Network*. Tabel 4.4 berikut menunjukkan hasil *confusion matrix* untuk *precision* dan *recall*

Tabel 4.3 Nilai *accuracy*, *precision*, *recall* ratio 70:30, *learning rate* 0.04

<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>
0.714	0.842	0.202

Selanjutnya, pada pengujian ini, *learning rate* yang diterapkan adalah sebesar 0.08, dan nilai *epoch* (jumlah iterasi selama pelatihan) sebanyak 50. *Learning rate* mempengaruhi seberapa besar perubahan yang akan diterapkan pada bobot (weights) model setiap kali pembaharuan dilakukan, sedangkan *epoch* menentukan seberapa sering seluruh dataset digunakan selama pelatihan. Pada gambar 4.4 menunjukkan hasil deteksi dengan parameter ratio data 70:30 dan *learning rate* 0.08.



Gambar 4.4 Confusion Matrix 70:30 pada *learning rate* 0.08

Hasil yang diberikan berhasil diprediksi dengan baik oleh *Artificial Neural Network*. Tabel 4.5 berikut menunjukkan hasil *confusion matrix* untuk *accuracy*, *precision* dan *recall*

Tabel 4.4 Nilai *accuracy*, *precision*, *recall* ratio 70:30, learning rate 0.08

<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>
0.766	0.842	0.202

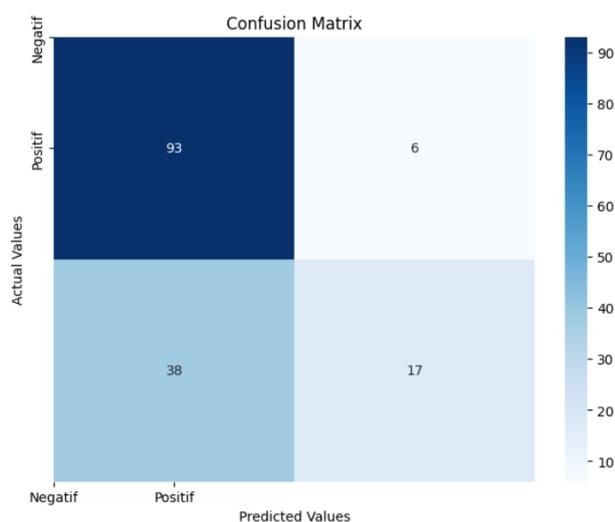
#### 4.2.2 Uji Coba Ratio Data 80:20

Dengan menerapkan skenario 1 menggunakan rasio 80:20 pada pengujian, digunakan learning rate sebesar 0,04 dan nilai *epoch* sebanyak 50. Penggunaan rasio 80:20 menunjukkan bahwa 80% dari data digunakan untuk pelatihan model, sementara 20% sisanya digunakan untuk pengujian. Rasio ini dapat mempengaruhi kinerja model dan kemampuannya untuk menggeneralisasi pada data yang tidak terlihat selama pelatihan.

*Learning rate* sebesar 0,04 digunakan untuk mengontrol sejauh mana model mengubah bobotnya setiap kali diperbarui selama pelatihan. Nilai ini dipilih dengan pertimbangan untuk mencapai keseimbangan antara konvergensi yang cepat dan menghindari overshooting. Selain itu, penggunaan 50 *epoch* menunjukkan bahwa model melalui 50 iterasi atau siklus melalui seluruh dataset pelatihan. Ini dapat memungkinkan model untuk memahami pola yang lebih kompleks dan meningkatkan kinerjanya seiring berjalannya waktu.

Dengan kombinasi parameter-parameter tersebut, skenario 1 diharapkan menghasilkan model yang mampu melakukan tugas yang diinginkan dengan tingkat akurasi dan generalisasi yang optimal. Namun, penting untuk terus memantau

kinerja model pada data uji dan mempertimbangkan penyesuaian parameter jika diperlukan untuk meningkatkan hasil secara keseluruhan. Sehingga diperoleh hasil dari pengujian tersebut menggunakan *confusion matrix* seperti pada gambar 4.5.



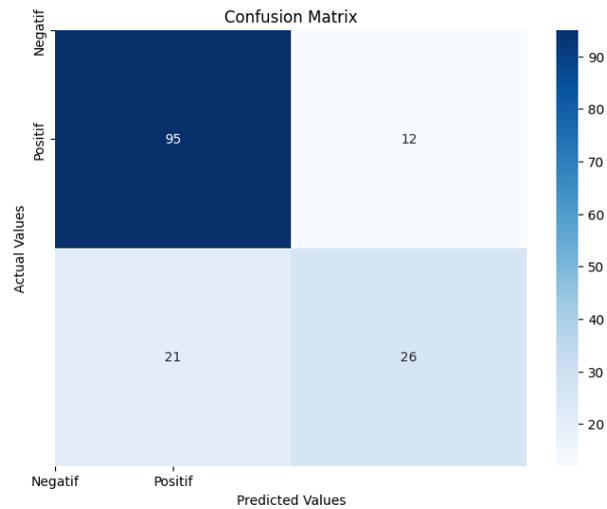
Gambar 4. 5 *Confusion Matrix* Hasil Deteksi

Gambar 4.5 menunjukkan hasil dari pengujian menggunakan *epoch* 50 dan *learning rate* 0.04. Hasil yang ditampilkan menunjukkan jumlah data yang berhasil dideteksi dengan baik menggunakan metode *Artificial Neural Network*. Berdasarkan perhitungan error tersebut, menghasilkan error sebesar 0.164.

Tabel 4.5 Nilai *accuracy*, *preccision*, *recall*

<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>
0.75	0.793	0.418

Pengujian selanjutnya dilakukan dengan mengubah parameter *learning rate* menjadi 0.08 dengan ratio data sama 80:20. Diperoleh hasil *confusion matrix* seperti pada gambar 4.6.



Gambar 4.6 Confusion Matrix 80:20 pada learning rate 0.08

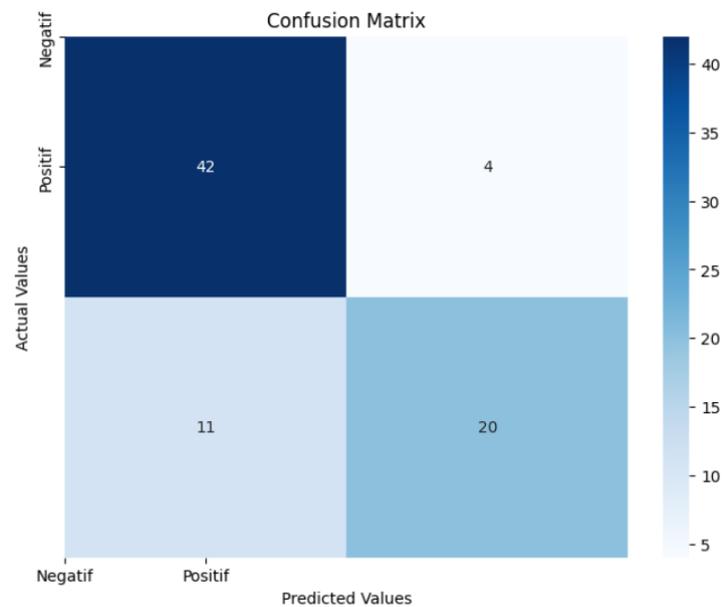
Berdasarkan Gambar 4.6 diperoleh nilai *accuracy*, *precision*, dan *recall* jika menggunakan parameter *learning rate* sebesar 0.08 yang dapat ditunjukkan pada Tabel 4.7.

Tabel 4. 6 Nilai *accuracy*, *precision*, *recall*

<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>
0.785	0.684	0.553

### 4.2.3 Hasil Uji Coba Ratio 90:10

Hasil uji coba pada skenario rasio 90:10 ini dilakukan dengan menerapkan *learning rate* sebesar 0.04 dan nilai *epoch* sebanyak 50. Kemudian menampilkan hasil dari pengujian tersebut menggunakan *confusion matrix* seperti pada gambar 4.7

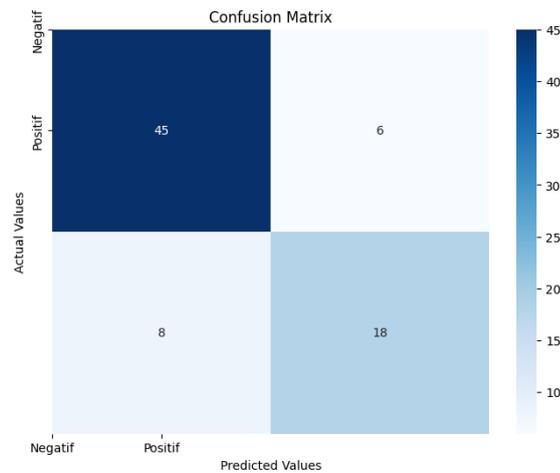


Gambar 4.7 Confusion Matrix 90:10, learning rate 0.04

Pada gambar 4.7 ini menunjukkan hasil dari pengujian skenario 3 yang menggunakan ratio 90:10, learning rate 0.04, dengan *epoch* 50. Hasil yang diberikan berhasil diprediksi dengan baik oleh *Artificial Neural Network*. Tabel 4.8 menunjukkan hasil *confusion matrix* untuk *precision* dan *recall*.

Tabel 4.7 Nilai *accuracy*, *precision*, *recall*

<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>
0.805	0.833	0.645



Gambar 4. 8 Confusion Matrix 90:10, *learning rate* 0.08

Pada gambar 4.8 ini menunjukkan hasil dari pengujian skenario yang menggunakan ratio 90:10, *learning rate* 0.08, dengan *epoch* 50. Hasil yang diberikan berhasil diprediksi dengan baik oleh *Artificial Neural Network*. Tabel 4.9 berikut menunjukkan hasil *confusion matrix* untuk *precision* dan *recall*

Tabel 4.8 Nilai Accuracy, Precision, Recall 90:10, *learning rate* 0.08

<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>
0.818	0.75	0.692

Hasil uji coba menggunakan ketiga ratio data yaitu 60:40, 70:30, dan 80:20 serta perubahan parameter *learning rate* 0.1, 0.2, 0.01, dan 0.03 pada epoch yang digunakan yaitu 50 mampu menghasilkan tingkat akurasi seperti pada Tabel 4.1.

Tabel 4. 9 Hasil Akurasi Seluruh Ratio

Ratio Data	Learning Rate	Akurasi
70:30	0.04	0.714
70:30	0.08	0.766
80:20	0.04	0.75
80:20	0.08	0.785
90:10	0.04	0.805
90:10	0.08	0.818

### 4.3 Pembahasan

Berdasarkan hasil uji coba yang dilakukan dengan membagi data keseluruhan menjadi 80% dari 768 data untuk data uji, dan 20% dari 768 data untuk data latih mampu menghasilkan akurasi terbaik dengan menggunakan *learning rate* yang sama yaitu 0.04.

Hasil uji coba menunjukkan bahwa terdapat aspek-aspek yang mempengaruhi terhadap nilai akurasi, presisi, dan *recall* tersebut. Berikut adalah aspek-aspek yang mempengaruhi nilai akurasi, presisi, dan *recall*:

#### a. Kualitas data

Kualitas data sangat mempengaruhi nilai akurasi, presisi, dan *recall* dalam konteks pemodelan dan evaluasi pembelajaran mesin. Data yang akurat, representatif, dan berkualitas tinggi akan membantu model mempelajari pola dengan baik, yang pada gilirannya dapat meningkatkan akurasi, presisi, dan *recall*. Kemudian aspek yang mempengaruhi yaitu jumlah data, jumlah sampel pelatihan juga memiliki pengaruh besar. Semakin banyak data yang digunakan untuk melatih model, semakin baik model dapat belajar pola dalam data tersebut.

#### b. *Preprocessing* data

Praproses data memiliki dampak besar pada nilai akurasi, presisi, dan *recall* dalam evaluasi model pembelajaran mesin. Praproses data melibatkan langkah-langkah seperti normalisasi data, penanganan data yang hilang, dan menangani kelas yang tidak seimbang. Sebagai mana pada gambar 4.3 yaitu normalisasi yang menghasilkan antara 0-1.

### c. Desain arsitektur ANN

Desain arsitektur *Artificial Neural Network* memengaruhi nilai akurasi, presisi, dan recall karena arsitektur ini menentukan bagaimana jaringan tersebut akan memproses data, mengekstrak fitur, dan mengambil keputusan. Termasuk jumlah lapisan dan neuron dalam setiap lapisan, dapat mempengaruhi performa model, pemilihan arsitektur yang tepat dapat meningkatkan akurasi.

Tujuan dilakukannya penelitian ini yaitu untuk mendeteksi diabetes ibu hamil berdasarkan fitur-fiturnya yang meliputi *pregnancies, glucose, blood pressure, skin thickness, insulin, BMI, diabetes pedigree function*, dan umur. Penelitian ini dilakukan agar penyakit diabetes yang apabila diderita pasien ini dapat di deteksi cepat supaya segera melakukan pengobatan agar nyawa pasien dapat segera diselamatkan. Sistem yang dibangun masih memerlukan beberapa perbaikan sehingga performa yang dihasilkan dapat lebih optimal.

Hal ini menunjukkan bagaimana kuasa Allah dalam menciptakan sesuatu dengan berbagai ukuran.

وَالْأَرْضَ مَدَدْنَاهَا وَالْقِيْنَ فِيهَا رَوَاسِي وَأَنْ بَتْنَا فِيهَا مِنْ كُلِّ شَيْءٍ مَّوْزُونٍ

“*Kami telah menghamparkan bumi, memancangkan padanya gunung-gunung, dan menumbuhkan di sana segala sesuatu menurut ukuran(-nya).*” (Q.S. *Al-Hijr*: 15-19)

Dalam ayat ini menunjukkan bahwa Allah telah menciptakan beraneka ragam tanam-tanaman dan tumbuh-tumbuhan, masing-masing mempunyai ukuran dan kadar yang ditentukan. Pohon durian yang batangnya kokoh itu serasi dengan buahnya yang besar dan berduri. Batang padi serasi dan sesuai pula dengan buahnya

yang bertangkai dan tanah untuk tempat tumbuhnya. Demikian pula tumbuh-tumbuhan yang lain diciptakan Allah seimbang, serasi, dan sesuai dengan iklim, keadaan daerah, dan keperluan manusia atau binatang tempat ia tumbuh. Sementara itu, perbedaan daerah dan tanah tempat tumbuh suatu pohon akan menimbulkan perbedaan rasa dan ukuran buahnya. Unsur gula di dalam tebu berlainan dengan unsur gula dalam air kelapa, berlainan manisnya dengan mangga dan jeruk. Buah salak sewaktu masih berupa putik dikelilingi oleh duri-duri yang tajam, tetapi setelah ia masak, seakan-akan duri-duri itu menguakkan diri, sehingga mudah untuk manusia mengambil buahnya yang rasanya manis. Putik pepaya pahit rasanya sewaktu masih kecil, sehingga manusia tidak mau mengambil dan memakannya. Semakin besar putiknya itu, semakin berkurang rasa pahitnya, dan semakin dekat pula manusia kepadanya. Setelah masak, buahnya dipetik dan menjadi makanan yang disenangi. Demikian Allah menciptakan sesuatu dengan ukuran dan kadar yang tertentu, sehingga melihat kesempurnaan ciptaan-Nya itu akan bertambah pula iman di dalam hati orang yang mau berpikir dan bertambah pula keyakinan bahwa Allah adalah Maha Sempurna.

Sistem pendeteksi penyakit diabetes ini dapat digunakan dalam bidang teknologi medis sebagai sarana untuk mencegah dan mengobati penyakit diabetes. Sehingga penyakit yang diderita mampu dideteksi dengan lebih cepat jika manusia mengalami gejala yang mengarah kepada penyakit diabetes atau penyakit lainnya dan bisa mendapatkan pertolongan dengan lebih cepat.

Dalam hal ini, QS. Luqman ayat 34:

انَّ اللّٰهَ عِنْدَهُ عِلْمُ السَّاعَةِ وَيُنزِلُ الْغَيْثَ وَيَعْلَمُ مَا فِي الْاَرْحَامِ قُلْ وَمَا تَدْرِي نَفْسٌ مَّاذَا تَكْسِبُ غَدًا قُلْ وَمَا تَدْرِي نَفْسٌ بِأَيِّ اَرْضٍ تَمُوتُ قُلْ اِنَّ اللّٰهَ عَلِيمٌ حَبِيرٌ (٣٤)

*“Sesungguhnya Allah memiliki pengetahuan tentang hari Kiamat, menurunkan hujan, dan mengetahui apa yang ada dalam rahim. Tidak ada seorang pun yang dapat mengetahui (dengan pasti) apa yang akan dia kerjakan besok. (Begitu pula,) tidak ada seorang pun yang dapat mengetahui di bumi mana dia akan mati. Sesungguhnya Allah Maha Mengetahui lagi Maha Teliti.” (Q.S. Luqman/ 31:34).*

Tafsir Jalalain pada ayat 34 Surah Luqman menunjukkan bahwa hanya Allah yang memiliki pengetahuan pasti mengenai tiga hal penting: kapan terjadi hari kiamat, kapan dan di mana hujan akan turun, dan jenis kelamin bayi yang akan lahir. Tidak ada yang dapat mengetahui dengan pasti apakah tindakan seseorang di masa depan akan menghasilkan hasil yang baik atau buruk. Ini mencerminkan bahwa masa depan tidak dapat diprediksi oleh manusia.

Penyakit diabetes ini disebabkan oleh pola gaya hidup yang modern dan diantaranya memilih gaya hidup yang praktis dalam berbagai hal, yakni dengan suka makan makanan cepat saji, kurangnya aktivitas fisik, karena lebih memilih menggunakan teknologi secara cepat. Pernyataan ini sesuai dengan surat al-baqarah ayat 168 sebagai berikut:

يَا أَيُّهَا النَّاسُ كُلُوا مِمَّا فِي الْأَرْضِ حَلَالًا طَيِّبًا وَلَا تَتَّبِعُوا خُطُوَاتِ الشَّيْطَانِ ۚ إِنَّهُ لَكُمْ عَدُوٌّ مُّبِينٌ

*“Hai sekalian manusia, makanlah yang halal lagi baik dari apa yang terdapat di bumi, dan janganlah kalian mengikuti langkah-langkah setan; karena sesungguhnya setan itu adalah musuh yang nyata bagimu” (QS Al-Baqarah : 168).*

Dari ayat Al-Qur'an Al-Baqarah ayat 168 dan hadits yang disampaikan oleh Muslim, Thabrani, Baihaqi, Ahmad, dan Abu Dawud, kita dapat merenungkan

bahwa sebagai seorang Muslim dan seorang profesional dalam bidang keperawatan, adalah wajib bagi kita untuk menjalani gaya hidup yang sehat. Seorang mukmin seharusnya hanya mengonsumsi makanan dan minuman yang telah dihalalkan, dan kita juga harus memperhatikan bahwa dalam aspek kesehatan, kita tidak boleh berlebihan sehingga makanan atau minuman yang dikonsumsi tidak akan berdampak negatif pada kondisi tubuh kita. Kita harus menyadari bahwa makanan dan minuman memiliki potensi untuk memberikan dampak buruk pada tubuh kita jika tidak diatur dengan baik.

## BAB V

### PENUTUP

#### 5.1 Kesimpulan

Berdasarkan hasil pengujian yang diperoleh, dapat disimpulkan bahwa *Artificial Neural Network* dapat digunakan dalam penelitian ini dengan beberapa uji coba dengan *learning rate* dan rasio yang berbeda memiliki tingkat akurasi yang besar. Pada penelitian ini menggunakan rasio sebesar 70%, 80%, dan 90% dari seluruh data yang berjumlah 768 data dengan *learning rate* 0.04 dan 0.08. Dari seluruh skenario ini akurasi terbesar terdapat pada rasio data 90:10 dengan *learning rate* 0.08 yaitu 0.818. Pada proses pengujian ini, *learning rate* mempengaruhi akurasi pada sistem, jika nilai *learning rate* semakin besar, maka hasil akurasi uji coba tersebut juga semakin meningkat.

#### 5.2 Saran

Saran yang dapat penulis sampaikan untuk penelitian selanjutnya adalah sebagai berikut:

- a. Dapat melakukan pengujian model dengan menambah model arsitektur *Artificial Neural Network*, *learning rate*, dan jumlah *epoch* yang lebih bervariasi untuk meningkatkan hasil akurasi prediksi.
- b. Dapat melakukan penelitian terkait dengan menggunakan metode *Machine Learning* lain untuk mendeteksi penyakit diabetes ibu hamil.

## DAFTAR PUSTAKA

- Abdi, P., Lubis, C., & Herwindiati, D. E. (2017). Deteksi Penyakit Diabetes Dengan Metode Fuzzy C-Means Clustering Dan K-Means Clustering. *Computatio : Journal of Computer Science and Information Systems*, 1(1), 15. <https://doi.org/10.24912/computatio.v1i1.233>
- Abedini, M., Bijari, A., & Banirostam, T. (2020). Classification of Pima Indian Diabetes Dataset using Ensemble of Decision Tree, Logistic Regression and Neural Network. *Ijarccce*, 9(7), 1–4. <https://doi.org/10.17148/ijarccce.2020.9701>
- Affonso, C., Sassi, R. J., & Barreiros, R. M. (2015). Biological image classification using rough-fuzzy artificial neural network. *Expert Systems with Applications*, 42(24), 9482–9488. <https://doi.org/10.1016/j.eswa.2015.07.075>
- Amato, F., López, A., Peña-méndez, E. M., Vañhara, P., & Hampl, A. (2013). *Artificial neural networks in medical diagnosis*. 47–58. <https://doi.org/10.2478/v10136-012-0031-x>
- Arhami, & Nasir. (2020). *DATA MINING Algoritma dan Implementasi*. ANDI.
- Budiyasari, V. N., Studi, P., Informatika, T., Teknik, F., Nusantara, U., & Kediri, P. (2017). Implementasi Data Mining Pada Penjualan kacamata Dengan Menggunakan Algoritma Apriori. *Indonesian Journal on Computer and Information Technology*, 2(2), 31–39.
- Carracher, A. M., Marathe, P. H., & Close, K. L. (2018). International Diabetes Federation 2017. *Journal of Diabetes*, 10(5), 353–356. <https://doi.org/10.1111/1753-0407.12644>
- Chang, V., Bailey, J., Xu, Q. A., & Sun, Z. (2022). Pima Indians diabetes mellitus classification based on machine learning (ML) algorithms. *Neural Computing and Applications*, 0123456789. <https://doi.org/10.1007/s00521-022-07049-z>
- Erlin, Yulvia Nora Marlim, Junadhi, Laili Suryati, & Nova Agustina. (2022). Deteksi Dini Penyakit Diabetes Menggunakan Machine Learning dengan Algoritma Logistic Regression. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, 11(2), 88–96. <https://doi.org/10.22146/jnteti.v11i2.3586>
- Gunawan, M. I., Sugiarto, D., & Mardianto, I. (2020). Peningkatan Kinerja Akurasi Prediksi Penyakit Diabetes Mellitus Menggunakan Metode Grid Search pada Algoritma Logistic Regression. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 6(3), 280. <https://doi.org/10.26418/jp.v6i3.40718>

- Hadi, H., & Ratnawati, E. (2018). ( NWKNN )Identifikasi Penyakit Gagal Ginjal Menggunakan Metode Neighbor Identifikasi Penyakit Gagal Ginjal Menggunakan Metode Neighbor Weighted K-Nearest Neighbor. *TEKNIMEDIA: Teknologi Informasi Dan Multimedia*, 8(2), 39–44.
- Irmawati, I., Widiyanto, K., Aziz, F., Rifai, A., & Rahmawati, A. (2022). Implementasi artificial neural network dalam mendeteksi penyakit hati (liver). *Journal of Information System, Applied, Management, Accounting and Research*, 6(1), 193–198. <https://doi.org/10.52362/jisamar.v6i1.69>
- Kamlesh, Lakhwani, S., Hiran, K. K., Bundele, M. M., & Somwanshi, D. (2020). Prediction of the Onset of Diabetes Using Artificial Neural Network and Pima Indians Diabetes Dataset. *2020 5th IEEE International Conference on Recent Advances and Innovations in Engineering, ICRAIE 2020 - Proceeding, 2020*. <https://doi.org/10.1109/ICRAIE51050.2020.9358308>
- Kurniawan, A. S. (2018). Implementasi Metode Artificial Neural Network Dalam Memprediksi Hasil Ujian Kompetensi Kebidanan (Studi Kasus Di Akademi Kebidanan Dehasen Bengkulu). *Pseudocode*, 5(1), 37–44. <https://doi.org/10.33369/pseudocode.5.1.37-44>
- Naz, H., & Ahuja, S. (2020). *Deep learning approach for diabetes prediction using PIMA Indian dataset*. 391–403.
- Niendy Alexandra Yosephine, & Ratnadewi. (2021). Penggunaan Artificial Neural Network pada Sinyal Elektrokardiogram untuk Mendeteksi Penyakit Jantung Aritmia Supraventrikular. *INFORMASI (Jurnal Informatika Dan Sistem Informasi)*, 13(1), 14–23. <https://doi.org/10.37424/informasi.v13i1.69>
- Nurhafifah Matondang, Mayanda Mega Santoni, N. C. (2019). *Deteksi Hipertensi Dengan Metode Artificial Neural Network Nurhafifah Matondang , Mayanda Mega Santoni , Nurul Chamidah*. 24–25.
- Nurul, Hamzah, Qomariyah, N., Prima Mustika, W., Studi Sistem Informasi, P., & Nusa Mandiri, S. (2018). Inti Nusa Mandiri Penerapan Metode Artificial Neural Network Untuk Mendeteksi Serangan Jantung Di Rs Awal Bros Bekasi. *Inti Nusa Mandiri*, 13(1), 27–32. Retrieved from [www.nusamandiri.ac.id](http://www.nusamandiri.ac.id)
- Pratama, A. P., Atmaja, R. D., & Tsp, H. F. (2016). *Deteksi Diabetes Retinopati Pada Foto Fundus Menggunakan Color Histogram & Transformasi Wavelet Diabetic Retinopathy Detection In Fundus Photographs Using Color Histogram & Wavelet Transform*. 3(3), 4552–4559.
- Purbolaksono, M. D., Irvan Tantowi, M., Imam Hidayat, A., & Adiwijaya, A. (2021). Perbandingan Support Vector Machine dan Modified Balanced

- Random Forest dalam Deteksi Pasien Penyakit Diabetes. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(2), 393–399. <https://doi.org/10.29207/resti.v5i2.3008>
- Purwanto, A., Darmadi, E. A., Trimitra, P., Mandiri, K., By, J., Jomin -Blok, P., & Barat -Kotabaru -Cikampek -Karawang, S.-J. (2018). *Perbandingan Minat Siswa Smu Pada Metode Klasifikasi Menggunakan 5 Algoritma*. 2(1), 43–47.
- Purwono, P., Dewi, P., Wibisono, S. K., & Dewa, B. P. (2022). Model Prediksi Otomatis Jenis Penyakit Hipertensi dengan Pemanfaatan Algoritma Machine Learning Artificial Neural Network. *Insect (Informatics and Security): Jurnal Teknik Informatika*, 7(2), 82–90. <https://doi.org/10.33506/insect.v7i2.1828>
- Pohan, S., Warsito, B., & Suryono, S. (2020). Backpropagation artificial neural network for prediction plant seedling growth. *Journal of Physics: Conference Series*, 1524(1). <https://doi.org/10.1088/1742-6596/1524/1/012147>
- Prajarini, D. (2022). Perbandingan Algoritma Klasifikasi Data Mining Untuk Prediksi Penyakit Stroke. *SINTECH (Science and Information Technology) Journal*, 5(2), 191–197. <https://doi.org/10.31598/sintechjournal.v5i2.1222>
- Rajni, & Amandeep. (2019). RB-bayes algorithm for the prediction of diabetic in “PIMA Indian dataset.” *International Journal of Electrical and Computer Engineering*, 9(6), 4866–4872. <https://doi.org/10.11591/ijece.v9i6.pp4866-4872>
- Riadi, A. (2017). Penerapan Metode Certainty Factor Untuk Sistem Pakar Diagnosa Penyakit Diabetes Melitus Pada Rsud Bumi Panua Kabupaten Pohuwato. *ILKOM Jurnal Ilmiah*, 9(3), 309–316. <https://doi.org/10.33096/ilkom.v9i3.162.309-316>
- Suwarno, A., Ferawati, N., Sari, P. A., Studi, P., Informatika, T., Teknik, F., ... Barang, S. (2021). Jurnal Teknologi Pelita Bangsa. *Jurnal Teknologi Pelita Bangsa*, 12(4), 33–40.
- Selano, M. K. (2021). Hubungan Lama Menderita Dengan Kejadian Neuropati Diabetikum Pada Pasien Diabetes Melitus. *Jurnal Smart Keperawatan*, 8(2), 129. <https://doi.org/10.34310/jskp.v8i2.505>
- Soltani, Z., & Jafarian, A. (2016). A New Artificial Neural Networks Approach for Diagnosing Diabetes Disease Type II. *International Journal of Advanced Computer Science and Applications*, 7(6), 89–94. <https://doi.org/10.14569/ijacsa.2016.070611>
- Syahrizal, M. (2018). *Perancangan Aplikasi Sistem Pakar Deteksi Kerusakan Mesin Alat Berat ( Beko ) Dengan Menerapkan Metode Teorema Bayes*. 2(2),

23–33.

Wafiah, A., Suwardoyo, U., & Syawal, F. (2021). *Assesorieshandphone Menggunakan Augmented. 1(1)*.

# **LAMPIRAN**

**Lampiran 1: Hasil Deteksi Ratio Data 70:30**

<i>Learning Rate</i>				
No	0.04		0.08	
	Aktual	Prediksi	Aktual	Prediksi
1	Yes	Yes	Yes	Yes
2	No	No	No	No
3	No	No	No	No
4	Yes	Yes	Yes	Yes
5	No	No	No	No
6	No	No	No	No
7	Yes	Yes	Yes	Yes
8	Yes	Yes	Yes	Yes
9	No	No	No	No
10	No	No	No	Yes
11	Yes	Yes	Yes	Yes
12	Yes	Yes	Yes	Yes
13	No	No	No	No
14	No	No	No	No
15	No	No	No	No
16	No	No	No	No
17	Yes	Yes	Yes	Yes
18	No	No	No	No
19	No	No	No	Yes
20	No	No	No	No
21	Yes	Yes	Yes	Yes
22	Yes	Yes	Yes	No
23	No	No	No	No
24	No	No	No	No
25	No	No	No	No
26	No	No	No	No
27	No	No	No	No
28	No	No	No	Yes
29	No	No	No	No
30	No	No	No	No
31	No	No	No	Yes
32	No	No	No	No
33	No	No	No	No
34	No	No	No	Yes
35	No	No	No	No
36	Yes	Yes	Yes	Yes
37	Yes	Yes	Yes	No
38	No	No	No	No
39	No	No	No	No

40	No	No	No	Yes
41	No	No	No	No
42	No	No	No	No
43	No	No	No	No
44	Yes	Yes	Yes	Yes
45	Yes	Yes	Yes	Yes
46	No	No	No	No
47	No	No	No	No
48	Yes	Yes	Yes	No
49	Yes	Yes	Yes	No
50	Yes	Yes	Yes	No
51	No	No	No	No
52	No	No	No	No
53	Yes	Yes	Yes	Yes
54	No	No	No	No
55	No	No	No	No
56	No	No	No	No
57	No	No	No	No
58	Yes	Yes	Yes	No
59	Yes	Yes	Yes	No
60	Yes	Yes	Yes	No
61	Yes	Yes	Yes	Yes
62	No	No	No	No
63	No	No	No	No
64	Yes	Yes	Yes	Yes
65	Yes	Yes	Yes	Yes
66	Yes	Yes	Yes	Yes
67	Yes	Yes	Yes	Yes
68	No	No	No	No
69	No	No	No	No
70	No	No	No	No
71	No	No	No	No
72	No	No	No	No
73	No	No	No	No
74	No	No	No	Yes
75	Yes	Yes	Yes	Yes
76	No	No	No	No
77	No	No	No	No
78	No	No	No	Yes
79	No	No	No	No
80	No	No	No	No
81	No	No	No	No
82	No	No	No	No

83	No	No	No	No
84	No	No	No	No
85	No	No	No	No
86	No	No	No	No
87	Yes	Yes	Yes	No
88	No	No	No	Yes
89	Yes	Yes	Yes	Yes
90	No	No	No	No
91	No	No	No	No
92	No	No	No	No
93	No	No	No	No
94	No	No	No	No
95	No	No	No	Yes
96	No	No	No	No
97	Yes	Yes	Yes	No
98	No	No	No	No
99	Yes	Yes	Yes	Yes
100	Yes	Yes	Yes	No
101	No	No	No	No
102	No	No	No	No
103	No	No	No	No
104	No	No	No	No
105	No	No	No	Yes
106	Yes	Yes	Yes	No
107	No	No	No	No
108	No	No	No	No
109	No	No	No	No
110	Yes	Yes	Yes	Yes
111	No	No	No	No
112	Yes	Yes	Yes	No
113	Yes	Yes	Yes	Yes
114	Yes	Yes	Yes	No
115	Yes	Yes	Yes	No
116	Yes	Yes	Yes	Yes
117	No	No	No	No
118	No	No	No	Yes
119	No	No	No	No
120	Yes	Yes	Yes	Yes
121	No	No	No	No
122	No	No	No	No
123	No	No	No	No
124	No	No	No	No
125	No	No	No	No

126	No	No	No	No
127	No	No	No	No
128	Yes	Yes	Yes	No
129	No	No	No	Yes
130	No	No	No	No
131	No	No	No	No
132	No	No	No	No
133	No	No	No	No
134	Yes	Yes	Yes	Yes
135	No	No	No	No
136	Yes	Yes	Yes	No
137	No	No	No	No
138	Yes	Yes	Yes	No
139	Yes	Yes	Yes	Yes
140	No	No	No	No
141	No	No	No	No
142	No	No	No	Yes
143	No	No	No	No
144	No	No	No	No
145	Yes	Yes	Yes	No
146	No	No	No	No
147	No	No	No	No
148	No	No	No	No
149	No	No	No	No
150	Yes	Yes	Yes	No
151	No	No	No	No
152	Yes	Yes	Yes	No
153	No	No	No	No
154	No	No	No	No
155	Yes	Yes	Yes	Yes
156	No	No	No	No
157	No	No	No	No
158	No	No	No	No
159	Yes	Yes	Yes	Yes
160	Yes	Yes	Yes	Yes
161	Yes	Yes	Yes	Yes
162	Yes	Yes	Yes	Yes
163	No	No	No	No
164	No	No	No	No
165	No	No	No	Yes
166	Yes	Yes	Yes	Yes
167	No	No	No	No
168	No	No	No	No

169	No	No	No	No
170	No	No	No	No
171	No	No	No	No
172	No	No	No	No
173	Yes	Yes	Yes	No
174	Yes	Yes	Yes	No
175	No	No	No	Yes
176	No	No	No	No
177	No	No	No	No
178	No	No	No	No
179	No	No	No	No
180	No	No	No	No
181	Yes	Yes	Yes	No
182	Yes	Yes	Yes	Yes
183	No	No	No	No
184	Yes	Yes	Yes	Yes
185	Yes	Yes	Yes	No
186	No	No	No	No
187	Yes	Yes	Yes	Yes
188	Yes	Yes	Yes	No
189	Yes	Yes	Yes	No
190	No	No	No	No
191	No	No	No	No
192	No	No	No	No
193	No	No	No	No
194	No	No	No	No
195	No	No	No	No
196	Yes	Yes	Yes	Yes
197	Yes	Yes	Yes	Yes
198	No	No	No	No
199	No	No	No	No
200	No	No	No	No
201	No	No	No	No
202	No	No	No	No
203	No	No	No	No
204	No	No	No	No
205	Yes	Yes	Yes	No
206	No	No	No	No
207	No	No	No	No
208	No	No	No	No
209	Yes	Yes	Yes	No
210	No	No	No	No
211	Yes	Yes	Yes	No

212	No	No	No	No
213	No	No	No	No
214	No	No	No	No
215	No	No	No	No
216	No	No	No	No
217	Yes	Yes	Yes	Yes
218	Yes	Yes	Yes	No
219	No	No	No	Yes
220	No	No	No	No
221	Yes	Yes	Yes	No
222	No	No	No	No
223	Yes	Yes	Yes	Yes
224	Yes	Yes	Yes	Yes
225	No	No	No	Yes
226	No	No	No	Yes
227	Yes	Yes	Yes	No
228	Yes	Yes	Yes	No
229	No	No	No	No
230	No	No	No	No
231	No	No	No	No

**Lampiran 2: Hasil Deteksi Ratio Data 80:20**

<i>Learning Rate</i>				
No	0.04		0.08	
	Aktual	Prediksi	Aktual	Prediksi
1	Yes	Yes	Yes	Yes
2	No	No	No	No
3	No	No	No	No
4	Yes	Yes	Yes	Yes
5	No	No	No	No
6	No	No	No	No
7	Yes	Yes	Yes	Yes
8	Yes	Yes	Yes	Yes
9	No	Yes	No	Yes
10	No	No	No	Yes
11	Yes	Yes	Yes	Yes
12	Yes	Yes	Yes	Yes
13	No	No	No	No
14	No	No	No	No
15	No	No	No	No
16	No	No	No	No
17	Yes	Yes	Yes	Yes
18	No	No	No	No

19	No	Yes	No	No
20	No	No	No	No
21	Yes	Yes	Yes	Yes
22	Yes	No	Yes	No
23	No	No	No	Yes
24	No	No	No	No
25	No	No	No	No
26	No	No	No	No
27	No	No	No	No
28	No	Yes	No	Yes
29	No	No	No	No
30	No	No	No	No
31	No	Yes	No	Yes
32	No	No	No	No
33	No	No	No	No
34	No	Yes	No	Yes
35	No	No	No	No
36	Yes	Yes	Yes	Yes
37	Yes	No	Yes	No
38	No	No	No	No
39	No	No	No	No
40	No	Yes	No	Yes
41	No	No	No	No
42	No	No	No	No
43	No	No	No	No
44	Yes	Yes	Yes	Yes
45	Yes	Yes	Yes	Yes
46	No	No	No	No
47	No	No	No	No
48	Yes	No	Yes	No
49	Yes	No	Yes	No
50	Yes	No	Yes	No
51	No	No	No	No
52	No	No	No	No
53	Yes	Yes	Yes	Yes
54	No	No	No	No
55	No	No	No	No
56	No	No	No	No
57	No	No	No	No
58	Yes	No	Yes	No
59	Yes	No	Yes	No
60	Yes	No	Yes	No
61	Yes	Yes	Yes	Yes

62	No	No	No	No
63	No	No	No	No
64	Yes	Yes	Yes	Yes
65	Yes	Yes	Yes	Yes
66	Yes	Yes	Yes	Yes
67	Yes	Yes	Yes	Yes
68	No	No	No	No
69	No	No	No	No
70	No	No	No	No
71	No	No	No	No
72	No	No	No	No
73	No	No	No	No
74	No	Yes	No	Yes
75	Yes	Yes	Yes	Yes
76	No	No	No	No
77	No	No	No	No
78	No	Yes	No	Yes
79	No	No	No	No
80	No	No	No	No
81	No	No	No	No
82	No	No	No	No
83	No	No	No	No
84	No	No	No	No
85	No	No	No	No
86	No	No	No	No
87	Yes	No	Yes	No
88	No	Yes	No	Yes
89	Yes	Yes	Yes	Yes
90	No	No	No	No
91	No	No	No	No
92	No	No	No	No
93	No	No	No	No
94	No	No	No	No
95	No	Yes	No	Yes
96	No	No	No	No
97	Yes	No	Yes	No
98	No	Yes	No	No
99	Yes	Yes	Yes	Yes
100	Yes	No	Yes	No
101	No	No	No	No
102	No	No	No	No
103	No	No	No	No
104	No	No	No	No

105	No	Yes	No	Yes
106	Yes	No	Yes	No
107	No	No	No	No
108	No	No	No	No
109	No	No	No	No
110	Yes	Yes	Yes	Yes
111	No	No	No	No
112	Yes	No	Yes	No
113	Yes	Yes	Yes	Yes
114	Yes	No	Yes	No
115	Yes	No	Yes	No
116	Yes	Yes	Yes	Yes
117	No	No	No	No
118	No	Yes	No	Yes
119	No	No	No	No
120	Yes	Yes	Yes	Yes
121	No	No	No	No
122	No	No	No	No
123	No	No	No	No
124	No	No	No	No
125	No	No	No	No
126	No	No	No	No
127	No	No	No	No
128	Yes	No	Yes	No
129	No	Yes	No	No
130	No	No	No	No
131	No	No	No	No
132	No	No	No	No
133	No	No	No	No
134	Yes	Yes	Yes	Yes
135	No	No	No	No
136	Yes	No	Yes	No
137	No	No	No	No
138	Yes	No	Yes	No
139	Yes	Yes	Yes	Yes
140	No	No	No	No
141	No	No	No	No
142	No	Yes	No	Yes
143	No	No	No	No
144	No	No	No	No
145	Yes	No	Yes	Yes
146	No	No	No	No
147	No	No	No	No

148	No	No	No	No
149	No	No	No	No
150	Yes	No	Yes	No
151	No	No	No	No
152	Yes	Yes	Yes	No
153	No	No	No	No
154	No	No	No	No

**Lampiran 3: Hasil Deteksi Ratio Data 90:10**

<i>Learning Rate</i>				
No	0.04		0.08	
	Aktual	Prediksi	Aktual	Prediksi
1	Yes	Yes	Yes	Yes
2	No	No	No	No
3	No	No	No	No
4	Yes	Yes	Yes	Yes
5	No	No	No	No
6	No	No	No	No
7	Yes	Yes	Yes	Yes
8	Yes	Yes	Yes	Yes
9	No	Yes	No	Yes
10	No	No	No	No
11	Yes	Yes	Yes	Yes
12	Yes	Yes	Yes	Yes
13	No	No	No	No
14	No	No	No	No
15	No	No	No	No
16	No	No	No	No
17	Yes	Yes	Yes	Yes
18	No	No	No	No
19	No	No	No	No
20	No	No	No	No
21	Yes	Yes	Yes	Yes
22	Yes	No	Yes	No
23	No	No	No	No
24	No	No	No	No
25	No	No	No	No
26	No	No	No	No
27	No	No	No	No
28	No	Yes	No	Yes
29	No	No	No	No

30	No	No	No	No
31	No	Yes	No	Yes
32	No	No	No	No
33	No	No	No	No
34	No	No	No	Yes
35	No	No	No	No
36	Yes	Yes	Yes	Yes
37	Yes	No	Yes	No
38	No	No	No	No
39	No	No	No	No
40	No	Yes	No	Yes
41	No	No	No	No
42	No	No	No	No
43	No	No	No	No
44	Yes	Yes	Yes	Yes
45	Yes	Yes	Yes	Yes
46	No	No	No	No
47	No	No	No	No
48	Yes	No	Yes	No
49	Yes	No	Yes	No
50	Yes	No	Yes	No
51	No	No	No	No
52	No	No	No	No
53	Yes	Yes	Yes	Yes
54	No	No	No	No
55	No	No	No	No
56	No	No	No	No
57	No	No	No	No
58	Yes	No	Yes	No
59	Yes	No	Yes	No
60	Yes	No	Yes	No
61	Yes	No	Yes	Yes
62	No	No	No	No
63	No	No	No	No
64	Yes	Yes	Yes	Yes
65	Yes	Yes	Yes	Yes
66	Yes	Yes	Yes	Yes
67	Yes	Yes	Yes	Yes
68	No	No	No	No
69	No	No	No	No
70	No	No	No	No
71	No	No	No	No
72	No	Yes	No	No

73	No	No	No	No
74	No	Yes	No	Yes
75	Yes	Yes	Yes	Yes
76	No	No	No	No
77	No	No	No	No
78	Yes	Yes	Yes	Yes