

**PENERAPAN *NEURAL NETWORK* DENGAN OPTIMASI *ANT COLONY OPTIMIZATION* DAN *BACKPROPAGATION* UNTUK MEMBANGUN MODEL PREDIKSI DIABETES TAHAP AWAL**

**SKRIPSI**

**Oleh:  
PYRENA HASNA ARTANTI  
NIM. 19650090**



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

**PENERAPAN *NEURAL NETWORK* DENGAN OPTIMASI *ANT COLONY OPTIMIZATION* DAN *BACKPROPAGATION* UNTUK MEMBANGUN MODEL PREDIKSI DIABETES TAHAP AWAL**

**SKRIPSI**

**Oleh:  
PYRENA HASNA ARTANTI  
NIM. 19650090**

**Diajukan kepada:  
Universitas Islam Negeri (UIN) Maulana Malik Ibrahim 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**

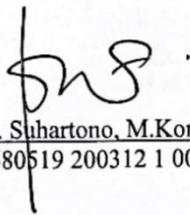
**PENERAPAN *NEURAL NETWORK* DENGAN OPTIMASI *ANT COLONY OPTIMIZATION* DAN *BACKPROPAGATION* UNTUK MEMBANGUN MODEL PREDIKSI DIABETES TAHAP AWAL**

**SKRIPSI**

**Oleh:  
PYRENA HASNA ARTANTI  
NIM. 19650090**

Telah Diperiksa dan Disetujui untuk Diuji:  
Tanggal: 21 Desember 2023

Pembimbing I,



Prof. Dr. Suhartono, M.Kom  
NIP. 19680519 200312 1 001

Pembimbing II,



Dr. Totok Chamidy, M.Kom  
NIP. 19691212 200604 1 001

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



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

## HALAMAN PENGESAHAN

### PENERAPAN *NEURAL NETWORK* DENGAN OPTIMASI *ANT COLONY OPTIMIZATION* DAN *BACKPROPAGATION* UNTUK MEMBANGUN MODEL PREDIKSI DIABETES TAHAP AWAL

#### SKRIPSI

Oleh:  
**PYRENA HASNA ARTANTI**  
NIM. 19650090

Telah Dipertahankan di Depan Dewan Penguji Skripsi  
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan  
Untuk Memperoleh Gelar Sarjana Komputer ( S.Kom )  
Tanggal: 21 Desember 2023

#### Susunan Dewan Penguji

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

Anggota Penguji I : Syahiduz Zaman, M.Kom  
NIP. 19700502 200501 1 005

Anggota Penguji II : Prof. Dr. Suhartono, M. Kom  
NIP. 19680519 200312 1 001

Anggota Penguji III : Dr. Totok Chamidy, M. Kom  
NIP. 19691222 200604 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, IPM  
NIP. 19771020 200912 1 001

## PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Pyrena Hasna Artanti

NIM : 19650090

Fakultas / Prodi : Sains dan Teknologi / Teknik Informatika

Judul Skripsi : Penerapan *Neural Network* Dengan Optimasi *Ant Colony Optimization* dan *Backpropagation* untuk Membangun Model Prediksi Diabetes Tahap Awal

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, 21 Desember 2023  
Yang membuat pernyataan,



Pyrena Hasna Artanti  
NIM.19650090

## **MOTTO**

*“Jika ingin menikmati manfaat dunia, juga harus menerima cobaannya.”*

*- Fang Haishi (Novoland: Pearl Eclipse) -*

## HALAMAN PERSEMBAHAN

Dengan rasa terimakasih yang sangat mendalam, saya panjatkan puji syukur dan terima kasih ke hadirat Allah SWT, yang dengan rahmat dan hidayah-Nya telah memberikan bimbingan dan perlindungan sehingga saya dapat menyelesaikan penulisan skripsi ini dengan sebaik-baiknya. Saya juga mengucapkan salawat serta salam kepada Nabi Muhammad SAW, yang telah mengajarkan kita kepada jalan kebenaran melalui ajaran Islam. Skripsi ini saya persembahkan juga sebagai bukti rasa terima kasih saya kepada semua pihak yang telah memberikan dukungan dan semangat dalam proses pembuatannya.

Skripsi ini saya persembahkan dengan segala hormat dan rasa terima kasih yang tak terhingga kepada kedua orang tua saya, Bapak Jemmy Sorongan dan Ibu Nuri Suhandini, yang telah memberikan dukungan, motivasi, dan doa tak terhingga sepanjang hidup saya. Saya juga ingin mengucapkan terimakasih kepada saudari-saudari saya mba Zara dan adik saya Widya yang selalu memberikan semangat dan dukungan untuk saya. Perjuangan, pengorbanan, dukungan dan doa yang telah mereka berikan adalah inspirasi dan dorongan bagi saya untuk menyelesaikan studi ini dengan baik. Saya juga ingin mengucapkan terima kasih kepada keluarga besar, dosen pembimbing, teman-teman, dan semua pihak yang telah memberikan bantuan, baik secara langsung maupun tidak langsung, yang telah menyertai saya dalam setiap langkah penyelesaian skripsi ini. Semoga Allah SWT memberikan balasan yang setimpal atas segala jerih payah mereka.

## KATA PENGANTAR

Segala puji saya panjatkan ke hadirat Allah Subhanahu Wa Ta'ala, Tuhan Semesta Alam, yang dengan segala karunia, kebijaksanaan, dan kasih sayang-Nya telah memberikan kemampuan dan kekuatan kepada peneliti untuk menyelesaikan penelitian dengan judul "Penerapan Neural Network dengan Optimasi Ant Colony Optimization dan Backpropagation untuk Membangun Model Prediksi Diabetes Tahap Awal" ini dengan baik. Penelitian ini merupakan upaya peneliti dalam memberikan kontribusi yang dapat bermanfaat dalam deteksi dini diabetes, yang diharapkan dapat memberikan dampak positif bagi kesehatan masyarakat.

Peneliti menyadari, tanpa rahmat dan hidayah dari-Nya, serta tanpa dukungan dari semua pihak, penelitian ini tidak akan dapat peneliti tuntaskan. Semoga penelitian ini dapat menjadi salah satu ikhtiar dalam rangka mencari solusi atas permasalahan kesehatan yang ada dan menjadi amal jariyah bagi peneliti semua. Shalawat serta salam semoga tercurah kepada Nabi Muhammad Shallallahu 'Alaihi Wasallam, utusan Allah SWT yang diutus sebagai rahmat bagi seluruh alam.

Penelitian ini tidak terlepas dari bimbingan, dukungan, dan doa dari berbagai pihak. Oleh karena itu, dengan penuh rasa syukur, peneliti mengucapkan terima kasih kepada para dosen pembimbing, keluarga, saudara, teman-teman dan semua yang terlibat secara langsung maupun tidak langsung dalam proses penelitian ini.

Penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. M. Zainuddin, MA selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.

2. Prof. Dr. Hj. Sri Harini, M.Si selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Fachrul Kurniawan, M.MT, IPM selaku Ketua Prodi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Prof. Dr. Suhartono, M. Kom, dan Dr. Totok Chamidy, M. Kom selaku dosen pembimbing skripsi dan dewan penguji yang dengan penuh kesabaran dan perhatiannya telah membimbing peneliti dalam mengerjakan dan menyelesaikan skripsi ini. Peneliti sangat berterimakasih pada arahan, masukan, dan saran yang diberikan selama mengerjakan skripsi ini.
5. Seluruh civitas akademika Jurusan Teknik Informatika, khususnya kepada seluruh Dosen dan Staf Jurusan Teknik Informatika Fakultas Sains dan Teknologi UIN Maulana Malik Ibrahim Malang, yang telah memberikan ilmu, pengalaman, dan bimbingan yang berharga sepanjang masa perkuliahan peneliti. Keberadaan dan dedikasi mereka semua telah membentuk pondasi yang kuat bagi perkembangan akademis dan profesional peneliti.
6. Nia Faricha, S.Si selaku Admin Prodi Teknik Informatika yang telah banyak membantu dan membimbing peneliti selama melaksanakan seminar-seminar dan selama penulisan skripsi berlangsung.
7. Teman-teman seperjuangan Angkatan 2019 Jurusan Teknik Informatika, kakak-kakak senior yang selalu memberikan doa dan semangat kepada penulis.
8. Sahabat penulis, Wildah Nourina dan adik penulis Widya Dhana yang telah memberikan kontribusi bantuan kepada penulis dikala laptop penulis sudah

rusak dan tidak bisa diperbaiki, mereka yang telah bersedia meminjamkan laptop nya untuk penulis gunakan demi melanjutkan dan menyelesaikan penulisan skripsi ini.

9. Seluruh Dosen dan Jajaran Staf Jurusan Teknik Informatika Fakultas Sains dan Teknologi UIN Maulana Malik Ibrahim Malang yang memberikan ilmu pengetahuan serta pengalaman.
10. Semua pihak yang tidak dapat disebutkan satu per satu, yang telah memberikan kontribusi, saran, dan dukungan dalam perjalanan penulisan skripsi ini.

Akhir kata, penulis juga menyadari bahwa skripsi ini masih memiliki keterbatasan, kekurangan, dan jauh dari kata sempurna. Oleh karena itu, dengan kerendahan hati, peneliti mempersembahkan penelitian ini sebagai wujud dari perjalanan akademis di Universitas Islam Negeri Maulana Malik Ibrahim Malang. Semoga apa yang peneliti sajikan dalam penelitian ini dapat memberikan nilai tambah dan menjadi inspirasi bagi penelitian-penelitian selanjutnya. Penulis juga menerima kritik dan saran yang membangun untuk perbaikan di masa depan.

Malang, 20 Desember 2023

Penulis

## DAFTAR ISI

<b>HALAMAN PERSETUJUAN</b> .....	ii
<b>HALAMAN PENGESAHAN</b> .....	ii
<b>PERNYATAAN KEASLIAN TULISAN</b> .....	iii
<b>MOTTO</b> .....	v
<b>HALAMAN PERSEMBAHAN</b> .....	vi
<b>KATA PENGANTAR</b> .....	vii
<b>DAFTAR ISI</b> .....	x
<b>DAFTAR GAMBAR</b> .....	xii
<b>DAFTAR TABEL</b> .....	xii
<b>ABSTRAK</b> .....	xiv
<b>ABSTRACT</b> .....	xv
<b>المُلخَص</b> .....	xvi
<b>BAB I PENDAHULUAN</b> .....	1
1.1 Latar Belakang .....	1
1.2 Pernyataan Masalah .....	8
1.3 Tujuan Penelitian .....	8
1.4 Batasan Masalah .....	9
1.5 Manfaat Penelitian .....	10
<b>BAB II TINJAUAN PUSTAKA</b> .....	11
2.1 Penelitian Terkait .....	11
2.2 Diabetes dan Risiko Tahap Awal .....	15
2.3 <i>Early Stage Diabetes Risk Prediction Dataset</i> .....	16
2.4 <i>Ant Colony Optimization (ACO)</i> .....	17
2.5 Penerapan ACO untuk Optimasi Bobot <i>Artificial Neural Network</i> .....	20
2.6 <i>Artificial Neural Network</i> .....	20
2.7 <i>Backpropagation</i> .....	24
<b>BAB III METODE PENELITIAN</b> .....	26
3.1 Alur Penelitian .....	26
3.2 Desain Sistem .....	28
3.3 Identifikasi Masalah .....	33
3.4 Studi Literatur .....	33
3.5 Eksplorasi Data .....	33
3.6 <i>Pre-Processing</i> .....	38
3.6.1 Mengubah Tipe Data Non-Numerik menjadi Numerik ( <i>Encoding</i> ) .....	38
3.6.2 Penskalaan Data .....	39
3.6.3 <i>Split Data</i> .....	41
3.7 <i>Artificial Neural Network</i> .....	42
3.8 <i>Ant Colony Optimization</i> .....	43
3.9 <i>Backpropagation</i> .....	51
3.10 Akurasi Hasil .....	52
<b>BAB IV UJI COBA DAN PEMBAHASAN</b> .....	53
4.1 Skenario Uji Coba .....	53

4.2 Hasil Uji Coba.....	55
4.2.1 MODEL A.....	55
4.2.2 MODEL B.....	60
4.2.3 MODEL C.....	65
4.3 Pembahasan.....	70
4.4 Integrasi Penelitian dalam Islam .....	77
4.4.1 Menjaga Kesehatan.....	79
4.4.2 Pentingnya Pengobatan.....	81
<b>BAB V KESIMPULAN DAN SARAN .....</b>	<b>82</b>
5.1 Kesimpulan .....	82
5.2 Saran.....	83
<b>DAFTAR PUSTAKA.....</b>	<b>85</b>

## DAFTAR GAMBAR

Gambar 2.1 Multi Layer Perceptron .....	22
Gambar 2.2 Single Neuron.....	22
Gambar 3.1 Diagram PlantUML Alur Penelitian.....	26
Gambar 3.2 Flowchart Desain Sistem Program.....	28
Gambar 3.3 Flowchart Model NN dioptimasi ACO dan BP.....	30
Gambar 3.4 Arsitektur ANN yang digunakan pada Penelitian .....	42
Gambar 3.5 Diagram <i>Flowchart</i> Optimasi <i>ACO</i> .....	48
Gambar 3.6 <i>Confusion Matrix</i> .....	52
Gambar 4.1 <i>Confusion Matrix</i> Model A, NN tanpa optimasi ACO dan BP .....	55
Gambar 4.2 <i>Confusion Matrix</i> Model A, NN dioptimasi dengan ACO.....	56
Gambar 4.3 <i>Confusion Matrix</i> Model A, NN dioptimasi dengan BP .....	57
Gambar 4.4 <i>Confusion Matrix</i> Model A, NN dioptimasi dengan ACO dan BP ...	58
Gambar 4.5 Perbandingan <i>Confusion Matrix</i> 4 Pelatihan pada Model A.....	59
Gambar 4.6 <i>Confusion Matrix</i> Model B, NN tanpa optimasi ACO dan BP .....	60
Gambar 4.7 <i>Confusion Matrix</i> Model B, NN dioptimasi dengan ACO.....	61
Gambar 4.8 <i>Confusion Matrix</i> Model B, NN dioptimasi dengan BP .....	62
Gambar 4.9 <i>Confusion Matrix</i> Model B, NN dioptimasi dengan ACO dan BP ...	63
Gambar 4.10 Perbandingan <i>Confusion Matrix</i> 4 Pelatihan pada Model B.....	64
Gambar 4.11 <i>Confusion Matrix</i> Model C, NN tanpa optimasi ACO dan BP .....	65
Gambar 4.12 <i>Confusion Matrix</i> Model C, NN dioptimasi dengan ACO.....	66
Gambar 4.13 <i>Confusion Matrix</i> Model C, NN dioptimasi dengan BP .....	67
Gambar 4.14 <i>Confusion Matrix</i> Model C, NN dioptimasi dengan ACO dan BP .	68
Gambar 4.15 Perbandingan <i>Confusion Matrix</i> 4 Pelatihan pada Model C.....	69
Gambar 4.16 Hasil BPNN dan KNN penelitian Jaka Permadi dkk. ....	72

## DAFTAR TABEL

Tabel 3. 1 Atribut-atribut pada dataset .....	34
Tabel 3.2 Tabel dataset diabetes tahap awal atribut 1-8 .....	36
Tabel 3.3 Tabel <i>dataset</i> diabetes tahap awal atribut 9-17 .....	37
Tabel 3.4 Tabel Distribusi Dataset Kategori <i>Class</i> .....	37
Tabel 3.5 Tabel Distribusi Dataset Kategori <i>Gender</i> Terhadap <i>Class</i> .....	37
Tabel 3.6 Hasil <i>Encoding</i> , Ubah Tipe Data Non-Numerik menjadi Numerik .....	39
Tabel 3.7 Delapan <i>Machine Learning</i> Penelitian Özge Nur Ergün dkk.....	74
Tabel 4.1 Perbandingan 4 Pembagian <i>Split</i> Data Latih dan Data Uji .....	53
Tabel 4.2 Perbandingan <i>Confusion Matrix</i> dari 4 Pelatihan pada Model A.....	59
Tabel 4.3 Perbandingan <i>Confusion Matrix</i> dari 4 Pelatihan pada Model B.....	64
Tabel 4.4 Perbandingan <i>Confusion Matrix</i> dari 4 Pelatihan pada Model C.....	69
Tabel 4.5 Hasil Performa <i>Confusion Matrix</i> Seluruh Model .....	71
Tabel 4.6 Perbandingan Algoritma Klasifikasi pada Penelitian Apriliah.....	73
Tabel 4.7 Perbandingan dengan Peneliti Lain dari Dataset yang Sama.....	75

## ABSTRAK

Artanti, Pyrena Hasna. 2023. **Penerapan *Neural Network* Dengan Optimasi *Ant Colony Optimization* dan *Backpropagation* untuk Membangun Model Prediksi Diabetes Tahap Awal**. Skripsi. Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Prof. Dr. Suhartono, M. Kom (II) Dr. Totok Chamidy, M. Kom

*Kata kunci: Diabetes, Neural Network, Ant Colony Optimization, Backpropagation*

Prevalensi diabetes yang mendunia merupakan permasalahan kesehatan yang memiliki peningkatan cepat, yang mempengaruhi jutaan orang dan terus bertambah sebagaimana yang dimuat pada data Diabetes Atlas dari International Diabetes Federation. Prediksi diabetes tahap awal yang akurat sangatlah penting untuk penanganan dan pengobatan yang efektif. Penggunaan Neural Network di dalam diagnosa medis telah memberikan harapan, meskipun Backpropagation tradisional yang digunakan masih memiliki kemungkinan untuk terjebak dalam local optima. Penelitian ini memperkenalkan model Neural Network yang dioptimasi menggunakan Ant Colony Optimization (ACO) dan Backpropagation (BP) untuk meningkatkan akurasi prediksi. Digunakan dataset dari UCI Machine Learning Repository terhadap 520 data pasien dengan 17 fitur sebagai dasar untuk pelatihan dan evaluasi model dengan pembagian data yang bervariasi untuk pelatihan dan pengujiannya. Perbandingan dilakukan pada tiga rasio split data dan pelatihan empat model, NN tanpa optimasi ACO dan BP, NN dengan optimasi ACO, NN dengan optimasi BP, dan NN dengan kombinasi optimasi ACO dan BP. Model gabungan optimasi dengan ACO dan BP menunjukkan kinerja yang unggul dengan akurasi 98.72%, presisi, recall dan F-measure mencapai 99.02% pada pembagian data latih dan data uji sebesar 70%:30. Penerapan ACO bersama dengan BP untuk pelatihan NN secara signifikan lebih unggul daripada metode konvensional, menyediakan strategi efektif untuk prediksi diabetes tahap awal. Pendekatan sistematis ini telah menetapkan model yang kuat, menunjukkan bahwa alokasi data pelatihan sebesar 70% merupakan pilihan optimal dalam konteks dataset ini.

## ABSTRACT

Artanti, Pyrena Hasna. 2023. **Application of Neural Network with Ant Colony Optimization and Backpropagation for Developing an Early-Stage Diabetes Prediction Model.** Thesis. Department of Informatics Engineering, Faculty of Science and Technology, State Islamic University of Maulana Malik Ibrahim Malang. Counselor: (I) Prof. Dr. Suhartono, M. Kom (II) Dr. Totok Chamidy, M. Kom

The global prevalence of diabetes is a mounting health concern, affecting millions and escalating rapidly, as indicated by the International Diabetes Federation's Diabetes Atlas. Accurate early-stage diabetes prediction is vital for effective management and treatment. The integration of Neural Networks (NN) in medical diagnostics has shown promise, yet traditional backpropagation training is often hindered by local optima entrapment. This research introduces an optimized Neural Network model employing Ant Colony Optimization (ACO) alongside Backpropagation (BP) for enhanced predictive accuracy. A dataset from the UCI Machine Learning Repository, comprising 520 patient samples with 17 features, served as the basis for model training and evaluation, with varying splits for training and testing. Comparisons were made across three data split ratios and four training models: NN without optimization, NN with ACO, NN with BP, and NN with both ACO and BP. The hybrid ACO-BP model demonstrated superior performance, with an accuracy of 98.72%, precision of 99.02%, and identical recall and F-measure, particularly at a 70:30 training-to-testing data allocation. The application of ACO in conjunction with BP for NN training significantly outperforms conventional methods, providing an effective strategy for early-stage diabetes prediction. This systematic approach has established a robust model, indicating that a 70% training data allocation is optimal within this dataset context.

**Key words:** Diabetes, Neural Network, Ant Colony Optimization, Backpropagation

## الملخص

أرتاني، بيرينا حسنا. 2023. "تطبيق الشبكة العصبية مع تحسين بالخوارزمية الاستعمارية للنمل والانتشار الخلفي لبناء نموذج تنبؤ مبكر بمرض السكري". أطروحة بكالوريوس. قسم هندسة المعلومات، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الوطنية في مالانج. المشرفون: (الأول) البروفيسور الدكتور سوهارتونو، ماجستير الحاسوب (الثاني) الدكتور توتوك تشاميدي، ماجستير الحاسوب.

انتشار مرض السكري على مستوى العالم يمثل مشكلة صحية متزايدة بسرعة، تؤثر على ملايين الأشخاص وتستمر في الزيادة كما هو موثق في "Diabetes Atlas" الصادر عن "International Diabetes Federation". الكشف المبكر والدقيق لمرض السكري أمر بالغ الأهمية للإدارة والعلاج الفعال. استخدام الشبكات العصبية في التشخيص الطبي قدم بصيص أمل، على الرغم من أن التدريب التقليدي باستخدام الانتشار الخلفي يظل محفوفاً بخطر الوقوع في الحد الأمثل المحلي. تقدم هذه الدراسة نموذجاً محسناً للشبكة العصبية باستخدام تقنية تحسين مستعمرة النمل والانتشار الخلفي لزيادة دقة التنبؤ. تم استخدام مجموعة بيانات من "UCI Machine Learning Repository" تضم 520 سجلاً لمرضى مع 17 خاصية كأساس لتدريب النموذج وتقييمه، مع تقسيم متفاوت للبيانات للتدريب والاختبار. أجريت مقارنات بين ثلاثة أنساق لتقسيم البيانات وأربعة نماذج تدريب: شبكة عصبية بدون تحسين باستخدام تقنية تحسين مستعمرة النمل والانتشار الخلفي، وشبكة عصبية مع تحسين باستخدام تقنية تحسين مستعمرة النمل، وشبكة عصبية مع تحسين باستخدام الانتشار الخلفي، وشبكة عصبية مع تحسين باستخدام كلا التقنيتين معاً. أظهر النموذج المدمج مع تحسين باستخدام تقنية تحسين مستعمرة النمل والانتشار الخلفي أداءً متفوقاً، حيث بلغت نسبة الدقة 98.72٪، والدقة والاسترجاع وقياس الفاعلية 99.02٪ عند تقسيم البيانات بنسبة 70٪ للتدريب و30٪ للاختبار. إن تطبيق تقنية تحسين مستعمرة النمل بالتزامن مع الانتشار الخلفي لتدريب الشبكات العصبية يتفوق بشكل كبير على الطرق التقليدية، ويوفر استراتيجية فعالة للتنبؤ المبكر بمرض السكري. لقد أسس هذا النهج المنظم لنموذج قوي، مما يشير إلى أن تخصيص 70٪ من البيانات للتدريب يعد الخيار الأمثل في سياق هذه المجموعة من البيانات.

الكلمات المفتاحية: السكري، الشبكة العصبية، تحسين مستعمرة النمل، الانتشار الخلفي

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Konsumsi makanan olahan yang terus meningkat di masyarakat telah menimbulkan dampak negatif terhadap kesehatan masyarakat. Menurut Pagliai dkk., (2021) makanan olahan, yang mengalami perubahan signifikan dengan penambahan zat seperti garam, gula, dan pengawet, memiliki kaitannya dengan masalah kesehatan seperti obesitas, diabetes tipe 2, dan penyakit kardiovaskular. (Gibney, 2018) menjelaskan, bahwa Ultra-Processed Food (UPF) memiliki resiko terhadap kesehatan karena adanya kandungan gula tambahan, lemak, rendahnya kadar serat, dan tingginya energi, yang berkontribusi pada faktor risiko penyakit seperti kardiovaskular dan kardiometabolik, obesitas, dan diabetes.

Diabetes merupakan salah satu penyakit tidak menular yang telah mempengaruhi jutaan orang di seluruh dunia dan merupakan salah satu penyakit yang prevalensinya semakin meningkat di seluruh dunia (WHO et al., n.d.). Laporan Edisi ke-10 Atlas Diabetes IDF (International Diabetes Federation, 2021) menunjukkan peningkatan prevalensi diabetes secara global, menegaskan bahwa diabetes merupakan tantangan besar bagi kesehatan dan kesejahteraan individu, keluarga, serta masyarakat secara global. Pada tahun 2021, tercatat sekitar 537 juta orang dewasa (usia 20-79 tahun) yang berarti 1 dari 10 orang dewasa hidup dengan memiliki diabetes. Angka ini diperkirakan akan meningkat menjadi 643 juta pada tahun 2030 dan 783 juta pada tahun 2045. Lebih dari 3 dari 4 orang dewasa

memiliki diabetes berada di negara-negara berpendapatan rendah dan menengah. Diabetes berkontribusi terhadap 6,7 juta kematian pada tahun 2021, atau setara dengan 1 kematian pada setiap 5 detiknya. Biaya kesehatan yang diakibatkan oleh diabetes mencapai setidaknya 966 miliar USD, meningkat 316% selama 15 tahun terakhir. Terdapat 541 juta orang dewasa yang memiliki IGT atau pre-diabetes melitus.

Di Asia Tenggara, 1 dari 11 orang dewasa (90 juta) hidup dengan diabetes. Jumlah ini diperkirakan akan naik menjadi 113 juta pada tahun 2030 dan 151 juta pada tahun 2045. Lebih dari 1 dari 2 orang dewasa dengan diabetes tidak terdiagnosis. Pada tahun 2021, tercatat 747.000 kematian akibat diabetes dan pengeluaran untuk diabetes mencapai USD 10 miliar. (International Diabetes Federation, 2021)

Di Indonesia, terjadi peningkatan jumlah orang dengan diabetes, dari 5.654,3 ribu pada tahun 2000 menjadi perkiraan 28.569,9 ribu pada tahun 2045. Prevalensi diabetes yang disesuaikan dengan usia juga meningkat dari 5,1% pada tahun 2000 menjadi diperkirakan 11,7% pada tahun 2045. Pada tahun 2000, tercatat 14.341,9 ribu orang dengan diabetes yang tidak terdiagnosis, mencakup 73,7% dari total kasus diabetes. Kasus Toleransi Glukosa Terganggu (IGT) juga tercatat meningkat, dari 12.190,2 ribu pada tahun 2000 menjadi perkiraan 34.623,1 ribu pada tahun 2045. Prevalensi IGT yang disesuaikan dengan usia juga mengalami peningkatan seiring waktu. (*Indonesia Diabetes Report 2000 — 2045*, n.d.). Sementara itu, jumlah orang dengan IFG naik dari 4.009,1 ribu pada tahun 2000 menjadi perkiraan 5.699,2 ribu pada tahun 2045, dengan sedikit peningkatan pada

prevalensi yang disesuaikan dengan usia. Data menunjukkan 149.872,0 kematian akibat diabetes pada tahun 2000, dengan 6,5% kematian terkait diabetes terjadi pada orang di bawah usia 60 tahun. Untuk diabetes tipe 1 pada anak-anak, terjadi penurunan kasus baru pada anak-anak berusia 0-14 tahun, dari 1.5 ribu pada tahun 2000 menjadi 0.7 ribu pada tahun 2011. Prevalensi pada kelompok umur ini menurun, sementara terjadi peningkatan pada kelompok umur 0-19 tahun. Pengeluaran kesehatan terkait diabetes mengalami peningkatan yang signifikan, dari 6.302,4 juta USD pada tahun 2000 menjadi 8.403,9 juta USD pada tahun 2030. Data demografis mencakup informasi mengenai populasi dewasa total (usia 20-79 tahun) dan populasi anak-anak (0-14 tahun), memberikan konteks untuk memahami proporsi populasi yang terpengaruh oleh diabetes. Laporan tersebut juga mencantumkan persentase komplikasi spesifik terkait diabetes, termasuk komplikasi mikrovaskular seperti nefropati, retinopati, dan neuropati, serta komplikasi makrovaskular seperti penyakit arteri koroner, penyakit serebrovaskular, penyakit arteri perifer, dan gagal jantung (International Diabetes Federation, 2021).

Penyebab munculnya diabetes melitus diakibatkan oleh berbagai faktor seperti adanya kerentanan genetik, gaya hidup, dan faktor lingkungan. Menurut (Nugroho et al., 2020) perlu adanya upaya yang dilakukan oleh Pemerintah dan lembaga kesehatan lainnya untuk meminimalkan risiko diabetes mellitus. Upaya untuk mengurangi risiko penyakit ini dapat berupa kebijakan atau gerakan-gerakan besar terhadap ajakan hidup sehat. Selain itu, memaksimalkan fungsi pemantauan

kesehatan secara teratur dapat menjadi upaya yang murah dan mudah untuk mendeteksi tanda-tanda risiko diabetes mellitus pada semua orang.

Menurut (Qowiyyum & Pradana, 2021), Lingkungan yang baik memegang peran penting dalam mendukung kesehatan masyarakat, menciptakan fondasi yang kuat bagi berbagai peningkatan kesehatan. Di sisi lain, status gizi masyarakat telah menunjukkan peningkatan signifikan, yang dihasilkan melalui edukasi kesehatan komprehensif, pemberian imunisasi kepada balita, dan pemantauan pertumbuhan masyarakat secara rutin. Selain itu, upaya edukasi yang berkelanjutan, adanya pencegahan dini, dan pemantauan kondisi lingkungan telah berhasil mengurangi tingkat kesakitan maupun kematian. Ketiga aspek ini saling terkait dan saling mendukung dalam menciptakan komunitas yang lebih sehat dan kuat.

Sebagaimana dikisahkan dalam Musnad Imam Ahmad, dari Ziyad bin Ilaqah, dari Usamah bin Syuraik

كُنْتُ عِنْدَ النَّبِيِّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ، وَجَاءَتِ الْأَعْرَابُ، فَقَالَ: يَا رَسُولَ اللَّهِ، أَتَدَاوَى؟ فَقَالَ: نَعَمْ يَا عِبَادَ اللَّهِ، تَدَاوَوْا، فَإِنَّ اللَّهَ عَزَّ وَجَلَّ لَمْ يَضَعْ دَاءً إِلَّا وَضَعَ لَهُ شِفَاءً غَيْرَ دَاءٍ وَاحِدٍ. قَالُوا: مَا هُوَ؟ قَالَ: الْهَرَمُ

*“Aku pernah berada di samping Rasulullah, Lalu datanglah serombongan Arab Badui. Mereka bertanya, 'Wahai Rasulullah, bolehkah kami berobat?' Beliau menjawab, 'Iya, wahai para hamba Allah, berobatlah. Sebab, Allah tidaklah meletakkan sebuah penyakit melainkan meletakkan pula obatnya, kecuali satu penyakit.' Mereka bertanya, 'Penyakit apa itu?' Beliau menjawab, 'Penyakit tua.'” (HR Ahmad).*

Hadis ini mengajarkan kita untuk selalu mencari pengobatan dan menjaga kesehatan. Meskipun tidak secara spesifik menyebutkan diabetes ataupun penyakit lainnya, hadis ini mengingatkan kita tentang pentingnya mencari pengetahuan dan pengobatan yang efektif untuk mengatasi penyakit, termasuk diabetes.

Dalam konteks diabetes, kita dapat menerapkan prinsip ini dengan mencari informasi tentang pengelolaan diabetes, menjalani gaya hidup sehat, dan bekerja sama dengan tenaga kesehatan untuk merencanakan pengobatan yang tepat. Selalu ingat bahwa menjaga kesehatan adalah bagian dari tanggung jawab kita sebagai umat Muslim, dan kita harus selalu berusaha untuk menjaga keseimbangan antara aspek fisik, mental, dan spiritual dalam kehidupan kita.

*Dataset* Diabetes Tahap Awal adalah *dataset* yang akan digunakan pada penelitian ini berdasarkan variabel-variabel dengan gejala-gejala yang berhubungan pada penyakit diabetes, seperti munculnya rasa gatal pada tubuh, seringnya pasien mengalami buang air kecil, luka yang lama sembuhnya, dan variabel penentu penyakit diabetes lainnya. *Dataset* ini dapat digunakan untuk mengklasifikasikan pasien ke dalam dua kategori, yaitu mereka yang memiliki diabetes (*Positive*) dan mereka yang tidak memiliki diabetes (*Negative*). Pemilihan teknik klasifikasi yang digunakan dalam *dataset* ini penting karena mereka dapat membantu dalam mengidentifikasi individu-individu yang berisiko memiliki diabetes dan memungkinkan untuk melakukan intervensi dini.

Pada saat yang sama, kemajuan teknologi telah membuka peluang baru dalam bidang medis, termasuk prediksi dan manajemen diabetes. Spesifiknya, teknik *Machine Learning* dan *Artificial Intelligence* (AI) telah digunakan untuk analisis data kesehatan dan prediksi penyakit, termasuk diabetes. *Deep learning* telah memperoleh posisi penting dalam beberapa tahun terakhir dalam bidang pembelajaran mesin dan pengenalan pola. *Deep learning* telah memungkinkan pengembangan solusi berbasis data dalam informatika kesehatan dengan

memungkinkan generasi fitur otomatis yang mengurangi jumlah intervensi manusia dalam proses ini.. (Ravi et al., 2016).

Dalam sebuah penelitian yang dilakukan oleh (Alić et al., 2017), Artikel ini membandingkan penggunaan *Artificial Neural Networks* (ANN) dan *Bayesian Networks* (BN) untuk klasifikasi diabetes dan penyakit kardiovaskular (CVD). Meskipun tingkat akurasi individu tertinggi dicapai pada BN untuk kedua penyakit tersebut, namun rata-rata akurasi dari beberapa studi menunjukkan lebih tinggi jika menggunakan ANN. Untuk diabetes, akurasi rata-rata ANN adalah 87,29%, dengan BN sebesar 80,98%. Untuk CVD, akurasi rata-rata ANN adalah 89,38%, dengan BN sebesar 86,49%. Selain itu, ANN secara konsisten menunjukkan akurasi yang lebih tinggi pada banyak kasus dibandingkan dengan BN. Meskipun BN memiliki potensi untuk akurasi yang lebih tinggi, hasilnya menunjukkan deviasi yang lebih besar dari nilai rata-rata, yang menunjukkan kurangnya keandalan. Studi ini menyimpulkan bahwa ANN memiliki potensi yang lebih tinggi untuk mencapai hasil yang lebih baik dan lebih dapat diandalkan untuk klasifikasi diabetes dan CVD.

ANN telah berhasil digunakan dalam berbagai aplikasi, termasuk klasifikasi dan prediksi penyakit untuk menyederhanakan proses diagnosis dalam rutinitas sehari-hari dan menghindari diagnosis yang salah, metode kecerdasan buatan (khususnya diagnosis yang dibantu komputer dan jaringan saraf tiruan) dapat digunakan. Algoritma pembelajaran adaptif ini dapat menangani berbagai jenis data medis dan mengintegrasikannya kedalam output yang telah dikategorikan. (Amato

et al., 2013). Dalam konteks diabetes, NN dapat digunakan untuk mempelajari pola dalam data kesehatan dan meramalkan risiko diabetes tahap awal.

Penggunaan klasifikasi *Neural Network* dalam melakukan diagnosis di bidang medis sudah semakin populer, terutama terhadap prevalensi penyakit kronis seperti Diabetes. Dalam sebuah penelitian oleh (Karegowda et al., 2011), penelitian ini memperkenalkan model gabungan (*hybrid*) antara *Genetic Algorithm* (GA) dan *Back Propagation Network* (BPN) untuk mendiagnosis diabetes mellitus menggunakan data medis. GA digunakan untuk menginisialisasi dan mengoptimalkan bobot koneksi dari BPN, serta menghasilkan peningkatan akurasi klasifikasi. Kesimpulannya, model GA-BPN hibrida secara efektif meningkatkan akurasi klasifikasi.

Pelatihan menggunakan *backpropagation* memiliki satu kelemahan yang terlihat jelas yaitu adanya kecenderungannya terjebak dalam *local optima*, yang menghalanginya untuk menemukan *global optima* (Whitley et al., 1990). Untuk mengatasi masalah ini, diadopsi penggunaan *Artificial Neural Networks evolutioner* (EANNs). Jaringan-jaringan ini memanfaatkan algoritma optimasi global untuk menemukan kombinasi bobot koneksi yang berada di dekat *optimum global*. Berbeda dengan algoritma *gradient descent*, algoritma optimasi global memiliki kemungkinan kecil untuk terjebak dalam solusi optimum lokal, (Mavrovouniotis & Yang, 2013). Algoritma optimasi global yang berguna lainnya adalah *Ant Colony Optimization* (ACO), yang awalnya diperkenalkan untuk menangani masalah optimasi diskrit, (Dorigo & Stützle, 2004).

Konsep *Ant Colony Optimization* (ACO) pertama kali diperkenalkan oleh Marco Dorigo dan koleganya pada awal 1990-an., (Dorigo et al., 1996) dimana pada pengembangannya terinspirasi dari perilaku koloni semut. Semut, sebagai serangga sosial, memprioritaskan kelangsungan hidup koloninya daripada kelangsungan hidup individu. Aspek kunci yang menginspirasi ACO adalah perilaku mencari makan semut, khususnya, kemampuan mereka untuk menemukan jalur terpendek antara sumber makanan dengan sarang mereka. Dalam pencarian makanannya, semut menjelajahi area di sekitar sarang mereka secara acak, meninggalkan jejak feromon kimia di tanah saat mereka bergerak. Jejak feromon ini, dapat dideteksi oleh semut lain, mempengaruhi jalur yang mereka pilih, dengan preferensi untuk jalur yang ditandai dengan konsentrasi feromon yang kuat. Ketika semut menemukan sumber makanan, semut tersebut mengevaluasi kuantitas dan kualitas makanan, lalu membawa beberapa kembali ke sarang, meninggalkan jejak feromon yang bervariasi tergantung pada kualitas dan kuantitas makanan. Perilaku ini memandu semut lain ke sumber makanan (Blum, 2005).

## **1.2 Pernyataan Masalah**

Bagaimana performa model Neural Network yang dioptimasi dengan Ant Colony Optimization (ACO) dan Backpropagation (BP) dalam mengklasifikasi dataset diabetes tahap awal (Early stage diabetes risk prediction dataset)?

## **1.3 Tujuan Penelitian**

Tujuan penelitian ini adalah untuk mengetahui performa model *Neural Network* yang dioptimasi dengan *Ant Colony Optimization* (ACO) dan

*Backpropagation (BP)* dalam mengklasifikasi dataset diabetes tahap awal (*Early stage diabetes risk prediction dataset*).

#### **1.4 Batasan Masalah**

Berdasarkan permasalahan yang telah dirumuskan, berikut adalah batasan masalah yang telah ditetapkan peneliti:

- a. Digunakan *dataset* publik yang diakses secara *online* melalui *UCI Machine Learning Repository*. Dataset merupakan kuisisioner dari pasien rumah sakit di Bangladesh, dan memiliki 17 atribut informasi data 520 individu dari *dataset* prediksi diabetes tahap awal “*Early stage diabetes risk prediction*”
- b. Hasil penelitian mungkin tidak dapat digeneralisasi ke dalam populasi yang lebih luas dikarenakan dataset tidak memiliki representasi demografis yang cukup untuk digunakan terhadap populasi lain.
- c. Karena penelitian ini terfokus pada deteksi diabetes tahap awal, hasilnya mungkin tidak berlaku untuk kasus diabetes yang lebih lanjut atau jenis diabetes lainnya.
- d. Perhitungan komputasi dan metode klasifikasi menggunakan bahasa pemrograman *Python* dan *libraries*-nya dengan *Google Colab* sebagai platform utamanya.
- e. Penelitian ini terbatas oleh pemilihan parameter yang digunakan (misalnya seperti jumlah iterasi, learning rate pada Neural Network, jumlah semut, dan kadar feromon dalam ACO) yang dapat mempengaruhi performa dan keakuratan model.

- f. Penelitian ini menggunakan metrik tertentu seperti *loss* dan *accuracy* untuk mengevaluasi performa model. Pemilihan metrik lain bisa menghasilkan interpretasi yang berbeda terkait efektivitas model.

### 1.5 Manfaat Penelitian

Berdasarkan permasalahan yang telah dirumuskan, berikut adalah manfaat yang dapat diperoleh dari penelitian ini:

- a. Manfaat dari penelitian ini adalah untuk menyediakan model prediksi yang lebih akurat untuk melakukan prediksi diabetes tahap awal, memfasilitasi deteksi dini dan membantu dalam perencanaan strategi perawatan yang lebih tepat bagi individu berisiko.
- b. Hasil penelitian ini juga bisa menjadi referensi bagi penelitian berikutnya yang berfokus pada penerapan algoritma *Ant Colony Optimization* dan *Backpropagation* dalam *Neural Network* untuk kasus prediksi lainnya.

## BAB II

### TINJAUAN PUSTAKA

#### 2.1 Penelitian Terkait

(Olabanjo et al., 2023) menerapkan model *deep unsupervised machine learning* untuk deteksi awal diabetes menggunakan voting seleksi fitur ensambel (*voting ensemble feature selection*) dan *Deep Belief Neural Networks* (DBN). Dataset yang diambil menggunakan data pertanyaan-pertanyaan yang diajukan kepada pasien-pasien pra-diagnosa dari rumah sakit Sylhet Bangladesh. Mereka melakukan preproses data. Fitur-fitur data dikurangi menggunakan *ensemble feature selector* seperti *Chi Square*, *Mutual Information Gain*, dan *Variance Threshold*. Penelitian tersebut juga menunjukkan bahwa fitur-fitur seperti *sex*, *polyuria*, *polydipsia*, *sudden weight loss*, *weakness*, *polyphagia*, *muscle stiffness*, dan *alopecia* adalah indikator yang kuat dalam dataset tersebut. Sementara *age*, *itching* dan *obesity* memberikan kontribusi yang cukup signifikan pada status diabetes dari pasien. Percobaan yang dilakukan peneliti menunjukkan performa yang lebih baik jika menggunakan fitur-fitur dengan indikator yang kuat seperti 8 fitur yang telah disebutkan sebelumnya meskipun ke-16 fitur juga menunjukkan hasil yang baik. Model DBN kemudian dilakukan pra-pemrosesan dan disesuaikan untuk mendapatkan performa yang optimal. Model tersebut juga dibandingkan dengan model-model lainnya yang tidak memiliki *multiple hidden layer* seperti *Decision Tree*, *Random Forest*, *Logistic Regression*, *Support Vector Machine*, dan *K-Nearest Neighbors*. DBN yang dijalankan menghasilkan kinerja *F1-measurement*,

*precision*, dan *recall* mencapai 1.00, 0.92, dan 1.00. Peneliti menyimpulkan bahwa model DBN adalah alat yang berguna untuk *unsupervised* prediksi awal dari diabetes tipe 2.

Pada penelitian (Aprilia et al., 2021) dilakukan penelitian terhadap algoritma *machine learning* seperti *Support Vector Machine* (SVM), *Naive Bayes* (NB), dan *Random Forest* (RF) untuk mendeteksi diabetes tahap awal. Dataset yang mereka gunakan adalah dataset diabetes tahap awal pasien rumah sakit Sylhet Bangladesh yang datanya diambil dari *UCI Machine Learning Repository*. Alat yang digunakan di dalam penelitian tersebut untuk menerapkan ketiga algoritma tersebut adalah aplikasi WEKA. Performa ketiga algoritma tersebut diukur dengan mengetahui nilai presisi, akurasi, *F-measure*, dan *recall* dari fitur-fitur yang diklasifikasikan ke dalam kelas sebagai pengidap diabetes atau bukan. Hasil percobaan menentukan kecukupan sistem yang dirancang dengan akurasi yang dicapai sebesar 97.88%. Hasil ini diverifikasi menggunakan kurva *Receiver Operating Characteristic* (ROC) secara tepat dan sistematis yang mencapai 99.8%

Pada penelitian (Permadi et al., 2021) dilakukan klasifikasi terhadap dataset pasien diabetes tahap awal dari rumah sakit Sylhet, Bangladesh dengan menggunakan dan membandingkan metode *K-Nearest Neighbor* (KNN) dan *Backpropagation Neural Network* (BPNN). Hasil dari penelitian yang dilakukan menunjukkan bahwa BPNN lebih unggul dibandingkan KNN dalam memprediksi diabetes tahap awal dengan tingkat akurasi sebesar 90%, presisi sebesar 90%, dan *recall* sebesar 90%, dengan *learning rate*  $\alpha = 0.3$ ,  $\alpha = 0.4$ , atau  $\alpha = 0.5$  dan jumlah *node hidden* = 5 unit. Sementara untuk KNN didapatkan hasil akurasi sebesar

83.75%, presisi sebesar 85.5497%, dan *recall* sebesar 83.75%, dari pengaturan nilai  $K = 5$ ,  $K = 7$ , atau  $K = 9$ .

Penelitian yang dilakukan (ERGÜN & O.İLHAN, 2021) dengan menggunakan dataset diabetes tahap awal dengan 16 fitur dari 520 orang pasien diabetes dari rumah sakit Sylhet, Bangladesh untuk membandingkan delapan model metode machine learning. Delapan metode tersebut adalah Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), XGBoost, K-Nearest Neighbor (KNN), Naive Bayes (NB), Artificial Neural Network (ANN), dan Convolutional Neural Network (CNN). Hasil dari perbandingan metode-metode tersebut di evaluasi dari akurasi, presisi, recall dan F-score dengan mengaplikasikan skema cross validation dimana pelatihan dan pengujian set diatur berdasarkan 416 dari 104 sampel pada tiap fold-nya. Pada hasil pengujian, didapatkan nilai akurasi tertinggi ada pada CNN dimensi 1 dengan nilai akurasi 99.04%. Pada nilai precision, performa CNN mencapai 100% dan nilai recall tertinggi adalah 98.63%. F-score tertinggi juga didapatkan dari model CNN dimensi 1 dengan nilai 99.31%. Penelitian tersebut juga menggunakan XGBoost dan model Decision Tree untuk menginvestigasi informasi fitur-fitur dari dataset. Berdasarkan performa klasifikasi XGBoost dengan akurasi 97.89%, precision 97.80%, dan recall 97.77%, menunjukkan bahwa usia pasien (*age*) memiliki parameter penting paling krusial dalam klasifikasi pasien diabeten. Kebanyakan pasien diklasifikasi dengan benar sebagai pengidap diabetes atau bukan berdasarkan informasi usia dari dataset. Parameter penting berikutnya yang krusial kedua adalah polydipsia. Fitur yang paling terakhir yang dapat menjadi parameter penentu diabetes atau bukan adalah

weight loss dan muscle stiffness. Pada Decision Tree, fitur paling penting sebagai penentu diabetes atau bukan adalah Ployuria yang kemudian memiliki cabang Polydipsia dan Age.

(Gao, 2011) mempresentasikan metode perkiraan baru untuk sistem keuangan yang kompleks, khususnya pada pasar saham, yang berdasarkan jaringan saraf evolusi yang baru. Jaringan ini menggabungkan algoritma BP tradisional dan algoritma ACO kontinu imun yang baru diajukan. Metode baru ini divalidasi dengan menggunakan data pasar saham dari pasar Shanghai pada tahun 1996, dan menunjukkan hasil yang menjanjikan. Hasil penemuan menunjukkan bahwa algoritma baru ini dapat dengan efektif memprediksi tren pasar saham, sehingga memiliki potensi aplikasi dalam investasi saham dunia nyata dan manajemen pasar.

(Wang et al., 2010) membahas tentang pemanfaatan jaringan saraf koloni semut untuk mengevaluasi sistem *Enterprise Resource Planning* (ERP). Model ini menggabungkan *neural network* BP dan algoritma koloni semut (ACO) untuk menilai efektivitas berbagai subsistem dalam sistem ERP. Analisis empiris dilakukan pada sistem ERP perusahaan batu bara yang terdiri dari tujuh subsistem, mengungkapkan bahwa subsistem 6 perlu ditingkatkan. Studi ini menemukan bahwa jaringan saraf koloni semut mengungguli jaringan saraf BP dalam hal akurasi, kecepatan, dan konvergensi jaringan. Model ini menguntungkan karena mengurangi beban kerja evaluasi dan meningkatkan rasionalitas hasil evaluasi, serta menunjukkan bahwa algoritma ini dapat diterapkan secara luas dalam evaluasi sistem ERP.

## 2.2 Diabetes dan Risiko Tahap Awal

Diabetes mellitus, yang sering disebut hanya diabetes, adalah sekelompok penyakit metabolik dimana ada tingkat kadar gula yang tinggi dalam jangka waktu yang lama di dalam tubuh. Gejala-gejala utama diabetes termasuk sering buang air kecil (poliuria), rasa haus berlebih dari biasanya (polidipsia), dan merasa lelah secara konstan. Diabetes adalah penyakit kronis yang terjadi ketika pankreas tidak memproduksi cukup insulin atau ketika tubuh tidak dapat menggunakan insulin yang diproduksi secara efektif. Insulin adalah hormon yang mengatur metabolisme gula (glukosa) dalam darah (American Diabetes Association, 2014).

Ada beberapa faktor risiko diabetes, termasuk obesitas, kebiasaan merokok, diet yang tidak sehat, kurang berolahraga, dan faktor genetik (International Diabetes Federation, 2019). Selain itu, individu dari kelompok etnis tertentu, seperti orang Asia Tenggara, cenderung lebih rentan terhadap efek diabetogenik dari lemak tubuh berlebih. Gejala diabetes biasanya meliputi sering buang air kecil, merasa sangat haus, merasa sangat lapar (meski baru saja makan), kelelahan yang tidak biasa, penurunan berat badan yang tidak dapat dijelaskan, penglihatan kabur, luka yang sulit sembuh, dan sering mengalami infeksi (Mayo Clinic Staff, 2023).

Deteksi dini dan pengelolaan risiko sangat penting dalam menangani diabetes. Penyakit ini dapat menyebabkan komplikasi jangka panjang yang serius jika tidak didiagnosis dan diperhatikan dengan baik, seperti penyakit jantung, stroke, kerusakan saraf, gangguan ginjal, masalah penglihatan, dan bahkan amputasi anggota badan bawah (American Diabetes Association, 2020). Selain itu, manajemen risiko juga penting untuk mencegah perkembangan dari prediabetes

menjadi diabetes tipe 2. Dengan deteksi dini, pengobatan yang tepat dapat dilakukan lebih dini, dan individu dapat membuat perubahan gaya hidup yang diperlukan untuk mencegah atau memperlambat perkembangan penyakit tersebut.

### ***2.3 Early Stage Diabetes Risk Prediction Dataset***

Dataset ini, berjudul "Prediksi Diabetes Tahap Awal," mencakup data tanda dan gejala dari pasien yang baru saja mengidap diabetes atau yang mungkin akan mengidap diabetes. Informasi ini dikumpulkan melalui kuesioner langsung dari pasien di Rumah Sakit Diabetes Sylhet di Sylhet, Bangladesh, dan telah mendapatkan persetujuan dokter. Dataset ini terdiri dari total 16 fitur dan satu variabel target bernama 'kelas'. Fitur-fitur tersebut meliputi Usia (dalam rentang 20-65 tahun), Jenis Kelamin, Poliuria, Polidipsia, Penurunan Berat Badan Mendadak, Kelemahan, Polifagia, Keputihan Genital, Penglihatan Kabur, Gatal, Irritabilitas, Penyembuhan yang Tertunda, Paresis Parsial, Kekakuan Otot, Alopecia, dan Obesitas. Variabel 'kelas' sendiri bisa bernilai positif (menunjukkan individu tersebut mengidap diabetes) atau negatif (menunjukkan individu tersebut tidak mengidap diabetes). Seluruh dataset tidak memiliki nilai yang hilang. Bentuk data dalam dataset ini melibatkan 16 atribut yang digunakan untuk memprediksi hasil kelas diabetes. Semua atribut, kecuali Usia, memiliki data kategorikal dengan dua kemungkinan hasil. (Islam et al., 2020).

## 2.4 *Ant Colony Optimization (ACO)*

Algoritma *Ant Colony Optimization* (ACO) diperkenalkan pertama kali oleh *Marco Dorigo* dan koleganya pada awal tahun 1990-an (Dorigo et al., 1996). Inspirasi dari algoritma ini berasal dari perilaku koloni semut. Dalam pencarian makanan, semut merambah area sekitar sarang mereka secara acak, dan meninggalkan jejak feromon di tanah. Jejak feromon ini, yang dapat dideteksi semut lain, dan akan mempengaruhi jalur yang mereka pilih. Semakin kuat konsentrasi feromon, semakin disukai jalur tersebut.

Konsep dasar dari ACO adalah menggunakan 'semut buatan' untuk mencari solusi optimal dalam suatu ruang pencarian dengan melepaskan 'feromon buatan' yang berfungsi sebagai penanda jalan. Semut yang menemukan solusi yang baik akan meninggalkan lebih banyak feromon, sehingga semut lain lebih mungkin untuk mengikuti jalur tersebut, yang akhirnya membantu mereka menemukan solusi yang lebih baik (Dorigo & Stützle, 2004)

*Pheromone Update Rule* dan *Probability Transition Rule*, (Dorigo et al., 2007). *Pheromone Update Rule* adalah ketika setelah semua semut menyelesaikan perjalanan mereka, kemudian feromon-feromon pada setiap jalur diperbarui. Jumlah feromon yang ditinggalkan oleh semut buatan pada jalur tertentu sebanding dengan kualitas solusi yang dihasilkan. Dengan kata lain, solusi yang lebih baik menghasilkan lebih banyak feromon. Sementara *Probability Transition Rule* adalah ketika semut buatan memilih jalur berikutnya berdasarkan probabilitas yang dipengaruhi oleh konsentrasi feromon dan jarak. Jalur dengan konsentrasi feromon

yang lebih tinggi dan jarak yang lebih pendek memiliki probabilitas lebih tinggi untuk dipilih.

Pada cara kerja ACO, setiap semut memilih jalur berdasarkan probabilitas yang dipengaruhi oleh 'feromon', sebuah zat kimia yang menandakan jalur yang pernah dilalui oleh semut lain, serta informasi heuristik lainnya, seperti jarak. Faktor-faktor ini kemudian digabungkan dalam sebuah rumus yang dinamakan probabilitas, yang fungsinya memastikan bahwa setiap semut akan membuat keputusan yang terinformasi berdasarkan pengalaman semut sebelumnya serta karakteristik lingkungan sekitar.

Rumus Probabilitas Pemilihan Jalur  $i$  ke  $j$ :

$$P_{ij} = \frac{(\tau_{ij}^{\alpha}) (\eta_{ij}^{\beta})}{\sum (\tau_{ij}^{\alpha}) (\eta_{ij}^{\beta})} \quad (2.1)$$

$\tau_{ij}$  = Feromon  $i$  ke  $j$

$\eta_{ij}$  = Informasi jarak jalur  $i$  ke  $j$

$\alpha$  dan  $\beta$  = Parameter yang mengontrol pengaruh daya tarik pheromone

Selanjutnya, karena lingkungan yang bersifat dinamis, feromon menguap seiring waktu. Pada machine learning, hal ini terjadi untuk mencegah sistem terjebak pada solusi suboptimal karena 'overcommitment' pada jalur yang pertama kali ditemukan. Ini diwujudkan dalam rumus dengan adanya pengurangan periodik pada nilai feromon di semua jalur, mendorong semut untuk terus menjelajahi dan menemukan solusi alternatif yang mungkin lebih efisien.

Rumus Penguapan Feromon:

$$\tau_{ij} = (1 - \rho) \times \tau_{ij} + \Delta\tau_{ij} \quad (2.2)$$

$\tau_{ij}$  = Intensitas feromon pada jalur  $i$  ke  $j$

$\rho$  = Laju penguapan feromon (Koefisien evaporasi)

$\Delta\tau_{ij}$  = Total peningkatan pheromone oleh semut

Selanjutnya semut akan meninggalkan feromon mereka sendiri. Jalur yang terbukti memberikan hasil yang baik akan secara bertahap diperkuat dengan adanya penambahan feromon, yang mengarahkan semut lain untuk mengikuti jalur tersebut dan secara bersama-sama mengoptimalkannya. Hal ini terjadi melalui proses dimana semut yang menyelesaikan perjalanan akan meningkatkan feromon pada jalur yang telah mereka tempuh berdasarkan efisiensi jalur tersebut.

Rumus Penambahan Feromon:

$$\Delta\tau_{ij} = \frac{Q}{L_k} \quad (2.3)$$

$\Delta\tau_{ij}$  = Total peningkatan pheromone oleh semut

$Q$  = Konstanta

$L_k$  = Cost, misalnya total jarak atau jalur yang ditempuh

Algoritma ACO ini dirancang untuk menemukan solusi terbaik dengan memadukan eksplorasi berbagai pilihan solusi dan fokus pada jalur yang telah terbukti efektif. Dengan siklus perhitungan yang konstan, ACO memadukan pendekatan pencarian acak dan pembelajaran bersama, menciptakan metode kuat dan adaptif untuk menyelesaikan berbagai tantangan kompleks di dunia nyata.

## 2.5 Penerapan ACO untuk Optimasi Bobot *Artificial Neural Network*

*Ant Colony Optimization* (ACO) dan *Artificial Neural Networks* (ANNs) dapat digabungkan untuk menciptakan sebuah sistem yang mampu belajar dan mengadaptasi diri untuk mencapai solusi yang optimal. Dalam konteks ini, ACO digunakan untuk melakukan pencarian global pada ruang bobot ANN, sedangkan ANN digunakan untuk memodelkan fungsi yang dicari (Bianchi et al., 2009).

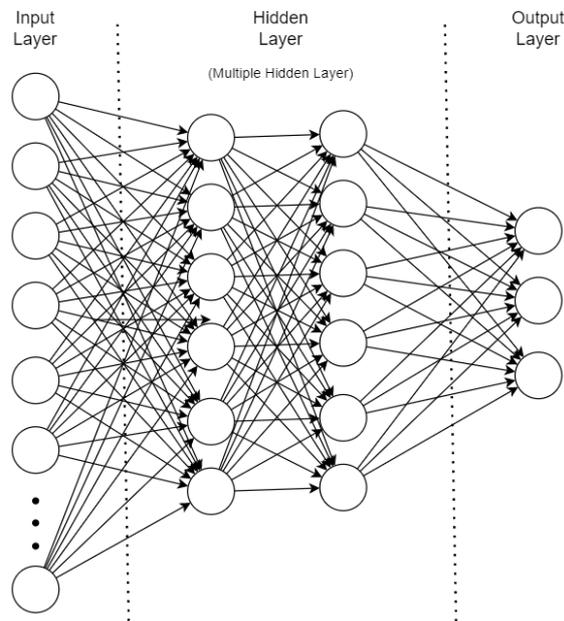
## 2.6 *Artificial Neural Network*

*Artificial Neural Network* adalah model komputasi yang terinspirasi oleh cara kerja sistem saraf manusia, terutama dalam hal pengolahan dan penyimpanan informasi. Arsitektur dasar ANN mencakup layer input, bobot dan bias, satu atau lebih dari satu layer tersembunyi (*hidden layers*), fungsi aktivasi, dan *layer output*. Setiap layer terdiri dari sejumlah unit atau '*neuron*' yang saling terhubung (Schmidhuber, 2015).

Seperti yang dijelaskan oleh (Goodfellow et al., 2016) fungsi aktivasi dalam ANN bertugas untuk menentukan apakah neuron harus "aktif" atau tidak berdasarkan informasi yang diterima. Ada banyak jenis fungsi aktivasi, seperti *step function*, *sigmoid*, *tanh*, dan *ReLU*. Sementara proses *forward pass* adalah proses di mana input data diteruskan melalui jaringan untuk menghasilkan *output*. Setelah *output* dihasilkan, kemudian kesalahan atau *error* antara *output* yang dihasilkan dan *output* yang diharapkan dihitung. Kesalahan ini kemudian digunakan dalam proses *backward pass*, di mana bobot dan bias diperbarui untuk meminimalkan kesalahan ini.

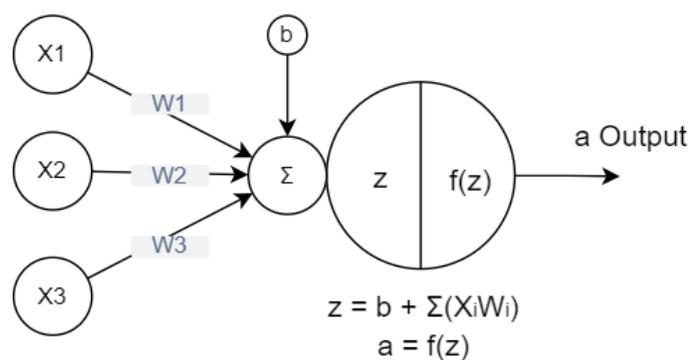
*Neural Network* dapat digunakan untuk melakukan tugas klasifikasi, seperti pengenalan pola dan prediksi. Contohnya, ANN dapat digunakan untuk mengklasifikasikan *email* sebagai *spam* atau bukan *spam*, atau untuk mengklasifikasikan gambar berdasarkan kategori tertentu (Zhang, 2000).

1. Menentukan arsitektur *Neural Network*: Pertama, tentukan berapa banyak jumlah *layer* (lapisan) dan *neuron* dalam setiap *layer* yang akan digunakan. Terdapat tiga jenis *layer*, yaitu *input*, *hidden* (tersembunyi), dan *output*. *Input layer* menerima data masukan, *hidden layer* melakukan pemrosesan, dan *output layer* menghasilkan prediksi atau keputusan. *Neural network* dengan satu lapisan *hidden layer* disebut sebagai "*shallow neural network*" atau "*single-layer hidden neural network*." Sedangkan *neural network* yang memiliki lebih dari satu *hidden layer* disebut sebagai "*deep neural network*" atau "*multi-layer hidden neural network*." *Deep neural networks* merupakan dasar dari *deep learning*, yang merupakan cabang dari *machine learning* yang fokus pada arsitektur *neural network* yang lebih kompleks dan dalam.



Gambar 2.1 Multi Layer Perceptron

2. Inisialisasi bobot dan bias: Bobot (*weight*) dan bias adalah parameter yang akan diubah selama proses pembelajaran untuk mengurangi kesalahan prediksi yang ada. Mereka diinisialisasi dengan nilai acak yang kecil.



Gambar 2.2 Single Neuron

3. *Feedforward*: Proses di mana data masuk melalui input melalui ANN dan menghasilkan output. Untuk setiap *neuron*, nilai input akan dikalikan dengan bobot yang sesuai dengan alur sinapsisnya, menjumlahkan hasil perkalian pada tiap *neuron/node*, kemudian menambahkan bias yang ada di *node* tersebut. Hasilnya kemudian dimasukkan ke dalam fungsi aktivasi (seperti *ReLU*, *sigmoid*, atau *tanh*) untuk menghasilkan *neuron output*.

$$z = \sum (x_i w_i) + b \quad (2.4)$$

$x_i = \text{input ke } - i$

$w_i = \text{bobot input ke } - i$

$b = \text{bias}$

$z = \text{total terbobot dari input, bobot, dan bias}$

$$a = f(z) \quad (2.5)$$

$f = \text{Fungsi aktivasi}$

$a = \text{aktivasi (output)}$

$z = \text{total terbobot dari input, bobot, dan bias}$

Secara Keseluruhan:

$$a = f\left(\sum (x_i w_i) + b\right) \quad (2.6)$$

Fungsi aktivasi adalah suatu fungsi matematika yang diaplikasikan pada output dari neuron dalam *artificial neural network* (ANN). Fungsi ini bertujuan

untuk mengubah sinyal input yang diterima oleh neuron menjadi sinyal output yang kemudian akan diteruskan ke neuron di layer berikutnya. Fungsi aktivasi juga menambahkan non-linearitas ke dalam ANN, yang memungkinkan ANN memodelkan suatu hubungan yang lebih kompleks antara input dan outputnya.

Fungsi aktivasi ReLU: ReLU adalah suatu fungsi aktivasi yang memperkenalkan non-linearitas ke dalam ANN dengan mengubah semua input negatif menjadi nol dan membiarkan input positif tidak berubah.

## **2.7 Backpropagation**

*Backpropagation*, atau propagasi balik, adalah algoritma pembelajaran yang digunakan dalam jaringan saraf tiruan (*Artificial Neural Network/ANN*). Algoritma ini bertujuan untuk meminimalkan kesalahan antara *output* aktual dan *output* yang diharapkan. Dengan kata lain, *backpropagation* digunakan untuk menyesuaikan bobot dan bias dalam ANN agar dapat menghasilkan *output* yang lebih akurat (Rumelhart et al., 1986)

Proses *backpropagation* melibatkan dua langkah utama: perhitungan gradien dan *update bobot*. Perhitungan gradien dilakukan dengan menggunakan turunan parsial dari fungsi kesalahan terhadap bobot dan bias. Setelah gradien dihitung, bobot dan bias kemudian diperbarui menggunakan metode *gradient descent* untuk meminimalkan fungsi kesalahan (Hecht-Nielsen, 1989).

Kelebihan utama dari *backpropagation* adalah kemampuannya untuk belajar dari kesalahan dan secara iteratif memperbarui bobot dan bias untuk mengurangi kesalahan tersebut. Sayangnya, *backpropagation* juga memiliki beberapa kekurangan, seperti adanya kemungkinan untuk terjebak dalam lokal

minima dan sensitivitas terhadap pilihan nilai awal bobot dan learning rate (Kadam, 2020).

Backpropagation menghitung gradien fungsi cost terhadap bobot dan bias menggunakan turunan parsial dan aturan rantai. Hal ini dilakukan secara mundur ke belakang dari output layer ke input layer.

Menggunakan aturan rantai :

$$\frac{\partial C}{\partial w} = \frac{\partial C}{\partial a} \frac{\partial a}{\partial z} \frac{\partial z}{\partial w} \quad \text{dan} \quad \frac{\partial C}{\partial b} = \frac{\partial C}{\partial a} \frac{\partial a}{\partial z} \frac{\partial z}{\partial b} \quad (2.7)$$

$\frac{\partial C}{\partial a}$  : Turunan fungsi biaya terhadap aktivasi

$\frac{\partial a}{\partial z}$  : Turunan fungsi aktivasi terhadap pre-aktivasi (tergantung pada fungsi aktivasi yang digunakan misalnya dari Sigmoid atau ReLU)

$\frac{\partial z}{\partial w}$  : Turunan pre-aktivasi terhadap bobot.

$\frac{\partial z}{\partial b}$  : Turunan pre-aktivasi terhadap bias.

Pembaruan Bobot dan Bias. Setelah kita menemukan gradien, gradien tersebut bisa digunakan untuk memperbarui bobot dan bias. Ini biasanya dilakukan dengan metode optimisasi seperti gradient descent. Pembaruan dilakukan dengan:

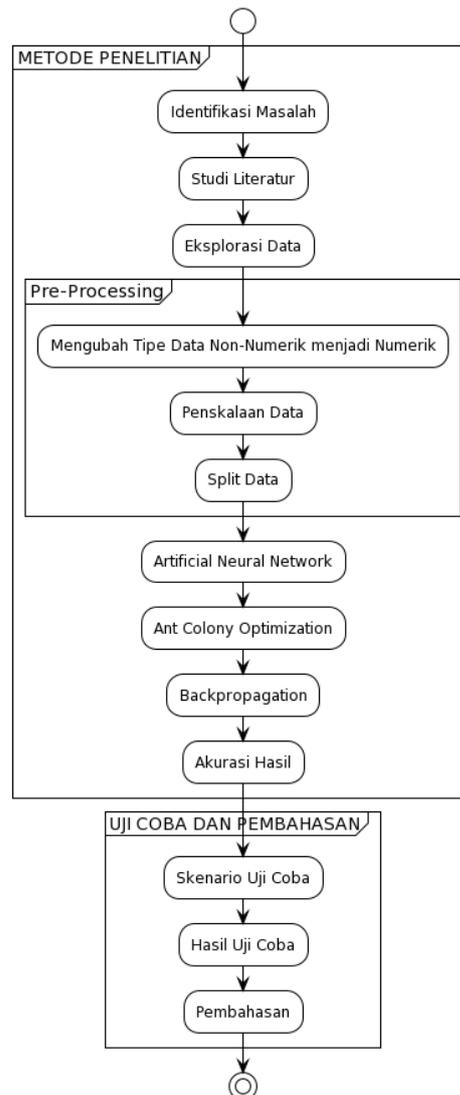
$$w = w - \eta \frac{\partial C}{\partial w} \quad \text{dan} \quad b = b - \eta \frac{\partial C}{\partial b} \quad (2.8)$$

di mana  $\eta$  adalah laju pembelajaran (*learning rate* nya).

## BAB III

### METODE PENELITIAN

#### 3.1 Alur Penelitian

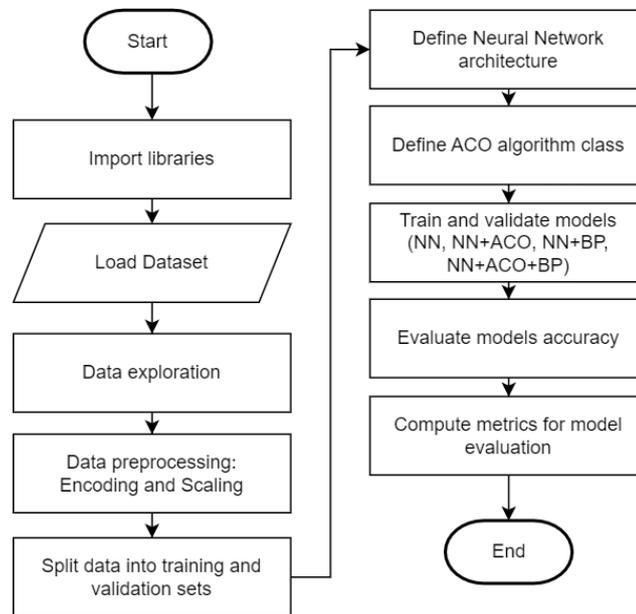


Gambar 3.1 Diagram PlantUML Alur Penelitian

Dalam penelitian ini, peneliti mengadopsi suatu alur sistematis untuk memastikan proses berjalan lancar dan efisien. Proses penelitian dimulai dengan identifikasi masalah, dimana masalah utama yang akan ditangani didefinisikan.

Setelah mengidentifikasi masalah, peneliti melanjutkan dengan Studi Literatur untuk memahami dan menganalisis teknik-teknik yang ada dan relevan dengan permasalahan yang dihadapi. Setelah memahami konteks, Eksplorasi Data dilakukan untuk memahami karakteristik dan struktur datanya. Langkah selanjutnya adalah Pra-pemrosesan Data, yang melibatkan penskalaan dan pengkodean komponen data yang diperlukan. Dengan data yang siap, peneliti kemudian melanjutkan dengan Implementasi Jaringan Saraf Tiruan untuk membangun struktur jaringan saraf tiruan yang optimal. Selanjutnya, untuk meningkatkan kinerja model, Optimasi menggunakan ACO diaplikasikan, dengan memanfaatkan Algoritma Ant Colony Optimization untuk penyetelan bobot. Proses pelatihan diperdalam lagi dengan tahapan Pelatihan menggunakan Backpropagation. Setelah model dilatih dengan baik, dilakukan Evaluasi Model untuk menilai performa model pada dataset validasi. Berdasarkan hasil evaluasi tersebut, Analisis Hasil dilakukan untuk mengambil kesimpulan dari metrik evaluasi. Penelitian berakhir setelah semua tahapan tersebut dilalui dengan teliti, menggambarkan urutan kegiatan yang dilakukan oleh peneliti dalam mengatasi masalah yang diajukan.

### 3.2 Desain Sistem



Gambar 3.2 Flowchart Desain Sistem Program

Langkah pertama adalah dengan mengimpor pustaka-pustaka Python yang diperlukan untuk melakukan operasi matematika, pengolahan data, penggunaan fungsi preprocessing, split data, framework deep learning, dan visualisasi data. Beberapa pustaka tersebut ada numpy, pandas, tensorflow, sklearn, seaborn, dan matplotlib.

Langkah berikutnya adalah memuat dataset. Karena dataset disimpan di dalam folder google drive, maka dilakukan mounting Google Drive dengan tujuan untuk menghubungkan Google Drive ke environment Colab agar dapat mengakses data yang disimpan di sana.

Selanjutnya dilakukan eksplorasi data. Eksplorasi data ini dilakukan untuk mengetahui beberapa informasi dari dataset seperti distribusi penyakit terhadap

gender, distribusi kasus positif dan negatifnya, mengetahui data mana yang memiliki nilai kosong, dan mengetahui tipe data dari fitur-fitur di dalam dataset.

Tahap selanjutnya adalah preprocessing yang meliputi encoding dan scaling fitur data. Fitur atau variabel yang masih kategorikal diubah ke dalam tipe numerik. Fitur yang memiliki rentang nilai yang besar, di skalakan agar memiliki rentang nilai yang mirip dengan fitur-fitur lainnya.

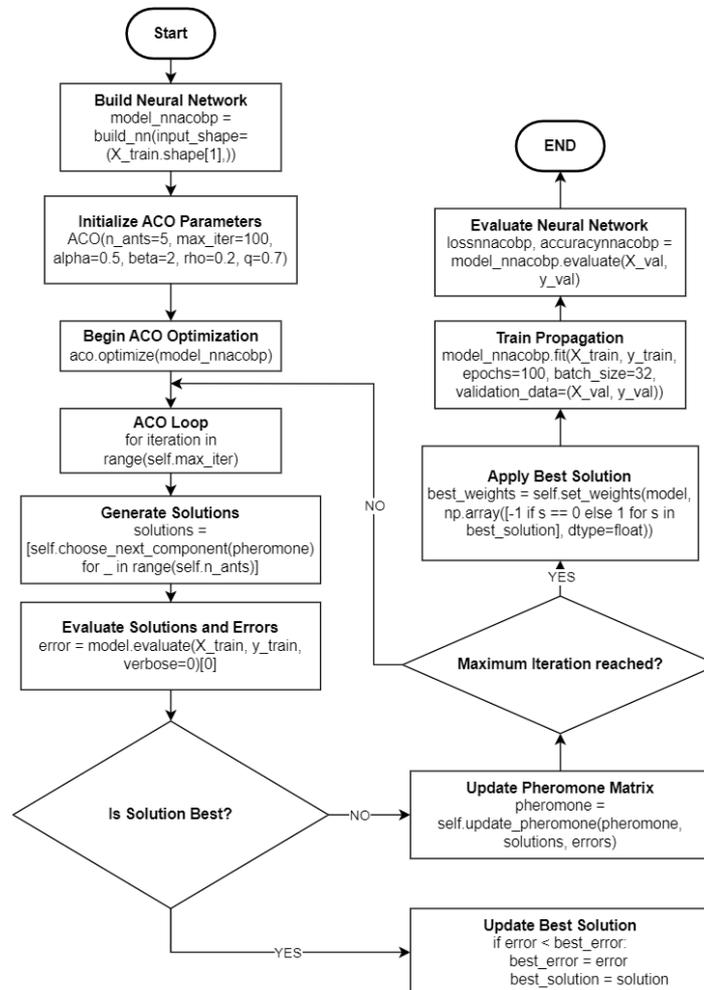
Berikutnya dilakukan split data dari dataset untuk membaginya ke dalam data yang akan dilatih dan data yang akan diuji.

Selanjutnya mendefinisikan arsitektur Neural Network yang akan digunakan untuk melatih dataset beserta parameter-parameternya seperti jumlah hidden layer, jumlah nodes, dan fungsi aktivasi apa yang digunakan pada tiap layer.

Selanjutnya mendefinisikan Class ACO yang berisi fungsi-fungsi untuk melakukan perhitungan algoritma semut. Class ACO ini mengadopsi prinsip algoritma semut pada fungsi-fungsinya dan akan digunakan untuk mengoptimalkan bobot pada jaringan neural network. Pada class ACO ini juga memiliki parameter-parameter seperti jumlah semut, maksimum iterasi, nilai alpha dan beta, serta nilai rho dan q.

Berikutnya dilakukan pelatihan dan pengujian model dengan empat skema pengujian yaitu pelatihan Neural Network tanpa optimasi dari ACO maupun BP, pelatihan Neural Network dengan optimasi menggunakan ACO, pelatihan Neural Network dengan optimasi menggunakan BP, dan pelatihan Neural Network dengan optimasi menggunakan ACO dan BP. Keempat pelatihan tersebut menggunakan parameter-parameter fixed.

Setelah pengujian dilakukan kemudian seluruh model dievaluasi hasil akurasi.



Gambar 3.3 Flowchart Model NN dioptimasi ACO dan BP

Gambar 3.3 adalah proses dari optimasi *Neural Network* menggunakan ACO dan BP

1. *Build Neural Network*: Proses ini melibatkan pembuatan arsitektur Artificial Neural Network dengan menggunakan 'Sequential' model dari pustaka Keras. Ini adalah langkah awal di mana struktur dasar NN dibuat.

Di sini, didefinisikan sebuah NN sederhana yang terdiri dari layer input, dua hidden layer, dan sebuah output layer.

2. *Initialize ACO Parameter*: Tahap ini berkaitan dengan menetapkan parameter-parameter untuk Ant Colony Optimization dalam kelas ACO yang telah dikembangkan. ACO adalah teknik optimasi yang terinspirasi dari perilaku koloni semut dalam mencari makanan.

3. *Begin ACO Optimization*: Pada tahap ini, proses optimasi ACO dimulai dengan menggunakan fungsi 'optimize' yang ada dalam kelas ACO. Fungsi ini bekerja dengan memanfaatkan ACO untuk mencari solusi optimal.

4. *Generate Solutions*: Dalam setiap iterasi solusi semut, dihasilkan solusi dan kemudian diubah menjadi set bobot yang akan digunakan dalam ANN. Ini memastikan bahwa solusi yang dihasilkan selalu diperbarui dan relevan.

5. *Evaluate Solutions and Errors*: Solusi yang telah dihasilkan dievaluasi terhadap model Neural Network berdasarkan bobot-bobot baru yang dihasilkan ini. Evaluasi ini membantu menentukan seberapa efektif solusi tersebut.

6. *Is Solution Best?* : Di sini, diperiksa apakah solusi yang ditemukan lebih baik daripada solusi sebelumnya. Ini adalah langkah kritis dalam menentukan efektivitas solusi yang dihasilkan.

7. *Update Best Solution*: Jika solusi yang dihasilkan terbukti lebih baik, maka solusi terbaik diperbarui. Ini membantu dalam menyimpan solusi terbaik yang ditemukan hingga saat itu.

8. *Update Pheromone Matrix*: Jika solusi yang dihasilkan bukan yang lebih baik, dan solusi terbaik telah ditetapkan, kemudian matrik feromon diperbarui. Pembaruan ini penting untuk memandu pencarian solusi yang lebih efektif.

9. *Maximum Iteration Reached?*: Di sini, diperiksa apakah jumlah maksimum iterasi yang telah ditentukan telah tercapai. Jika sudah, proses optimasi berhenti.

10. *Apply Best Solution*: Solusi terbaik yang telah ditemukan kemudian diterapkan sebagai bobot model dalam Neural Network. Ini merupakan implementasi dari solusi optimal yang ditemukan.

11. *Train Propagation*: Setelah menerapkan solusi terbaik, Neural Network kemudian dilatih dengan bobot baru ini menggunakan metode Backpropagation. Ini membantu dalam meningkatkan akurasi model.

12. *Evaluate Neural Network*: Akhirnya, Neural Network yang telah dilatih dievaluasi menggunakan data validasi. Ini memberikan wawasan tentang seberapa baik model tersebut bekerja pada data yang belum pernah dilihat sebelumnya.

terdapat proses optimasi dari model Neural Network dengan ACO (*Ant Colony Optimization*) dan BP (*Backpropagation*). Pada penelitian ini, terdapat 4 skema pelatihan. Pertama ada pelatihan dataset menggunakan NN tanpa optimasi dari ACO dan BP, NN dengan optimasi ACO, NN dengan optimasi BP, dan NN dengan optimasi ACO dan BP. Pelatihan 4 skema tersebut bertujuan

untuk membandingkan perbedaan Neural Network yang dioptimasi dengan ACO dan BP dengan tanpa optimasi dari ACO dan BP.

### **3.3 Identifikasi Masalah**

Di awal penelitian, langkah pertama yang dilakukan oleh peneliti adalah mengidentifikasi dan mendefinisikan masalah yang menjadi fokus penelitian. Identifikasi masalah ini penting untuk menentukan arah dan tujuan dari penelitian. Dengan memahami masalah dengan jelas, peneliti dapat merancang metodologi yang tepat dan menentukan teknik-teknik yang paling sesuai untuk mencapai tujuan penelitian.

### **3.4 Studi Literatur**

Setelah mengidentifikasi masalah, peneliti melakukan studi literatur yang mendalam. Tahap ini melibatkan kajian terhadap berbagai sumber literatur, baik berupa buku, jurnal penelitian, artikel, maupun website, yang berkaitan dengan topik penelitian. Dengan melakukan studi literatur, peneliti dapat memahami perkembangan terbaru dalam bidang yang diteliti, teknik-teknik yang telah ada, serta metodologi yang sering digunakan oleh peneliti lain. Studi literatur juga membantu dalam mengidentifikasi celah-celah dalam penelitian sebelumnya yang dapat diisi oleh penelitian ini.

### **3.5 Eksplorasi Data**

Penelitian yang akan dilakukan menggunakan, informasi berupa data yang akan diolah sebelum diproses lebih lanjut. Sumber data yang digunakan dalam penelitian ini diperoleh dari situs *UCI Machine Learning Repository*, yang

merupakan platform komunitas *Machine Learning* untuk memperoleh data secara publik. Data yang diambil bersifat terbuka untuk umum. Selanjutnya, data publik ini dibagi menjadi dua bagian penting, yaitu data latih dan data uji, yang memiliki peran penting untuk membangun model sistem yang akan dikembangkan.

Data latih dan data uji memiliki fungsi yang berbeda, namun keduanya saling mendukung dalam proses pengembangan model sistem. Data latih berfungsi sebagai fondasi dalam membangun model pelatihan yang akan digunakan untuk mengajarkan sistem dalam mengenali pola dan melakukan klasifikasi dari keputusan yang akan dihasilkan. Di sisi lain, data uji digunakan untuk menguji kinerja model yang telah dibangun dengan menggunakan data latih sebelumnya, agar dapat menilai sejauh mana sistem mampu bekerja dengan baik dan akurat. Dengan demikian, kedua jenis data ini menjadi unsur yang saling melengkapi dalam menciptakan sistem yang handal dan efisien.

Atribut-atribut informasi dari *dataset* tersebut memiliki atribut usia dengan rentang 20 sampai 65 tahun, jenis kelamin pria dan wanita, apakah individu mengalami gejala-gejala yang mengarah kepada penderita diabetes, apakah individu mengalami perubahan pada tubuhnya. Terdapat total 17 atribut dengan 520 data pasien. Atribut *dataset* yang digunakan adalah sebagai berikut:

Tabel 3. 1 Atribut-atribut pada dataset

No	Parameter	Keterangan
1	<b>Usia (<i>Age</i>)</b>	Usia pasien dengan rentang 20 sampai 65 tahun. Tipe data integer.
2	<b>Jenis kelamin (<i>Sex</i>)</b>	Jenis kelamin pasien wanita atau pria. Tipe data awal <i>object</i> , akan diubah setelahnya menjadi integer dimana 1 untuk pria ( <i>male</i> ), dan 0 untuk wanita ( <i>female</i> )

3	<b><i>Polyuria</i></b>	Apakah pasien mengalami peningkatan frekuensi buang air kecil (urinasi) dan volume urin yang lebih banyak dari normal. Tipe data awal <i>object</i> , akan diubah setelahnya menjadi integer dimana 1 untuk jawaban “yes” dan 0 untuk jawaban “no”
4	<b><i>Polydipsia</i></b>	Apakah pasien mengalami rasa haus yang berlebihan dan seringkali sulit dipuaskan. Tipe data awal <i>object</i> , akan diubah setelahnya menjadi integer dimana 1 untuk jawaban “yes” dan 0 untuk jawaban “no”
5	<b>Penurunan berat badan secara tiba-tiba. (<i>Sudden weight loss</i>)</b>	Apakah pasien mengalami penurunan berat badan secara tiba-tiba. Tipe data awal <i>object</i> , akan diubah setelahnya menjadi integer dimana 1 untuk jawaban “yes” dan 0 untuk jawaban “no”
6	<b>Merasa lemas (<i>Weakness</i>)</b>	Apakah pasien mengalami tubuh yang lemas. Tipe data awal <i>object</i> , akan diubah setelahnya menjadi integer dimana 1 untuk jawaban “yes” dan 0 untuk jawaban “no”
7	<b><i>Polyphagia</i></b>	Apakah pasien mengalami peningkatan rasa lapar atau nafsu makan yang berlebihan. Tipe data awal <i>object</i> , akan diubah setelahnya menjadi integer dimana 1 untuk jawaban “yes” dan 0 untuk jawaban “no”
8	<b>Adanya pertumbuhan jamur pada organ genital (<i>Genital thrush</i>)</b>	Apakah ada pertumbuhan jamur di organ genital pasien. Tipe data awal <i>object</i> , akan diubah setelahnya menjadi integer dimana 1 untuk jawaban “yes” dan 0 untuk jawaban “no”
9	<b>Pandangan yang kabur/buram (<i>Visual Blurring</i>)</b>	Apakah pandangan pasien mulai buram. Tipe data awal <i>object</i> , akan diubah setelahnya menjadi integer dimana 1 untuk jawaban “yes” dan 0 untuk jawaban “no”
10	<b>Gatal-gatal (<i>Itching</i>)</b>	Apakah pasien mengalami gatal-gatal. Tipe data awal <i>object</i> , akan diubah setelahnya menjadi integer dimana 1 untuk jawaban “yes” dan 0 untuk jawaban “no”

11	<b>Perasaan mudah tersinggung (<i>Irritability</i>)</b>	Apakah pasien mulai merasa mudah tersinggung. Tipe data awal <i>object</i> , akan diubah setelahnya menjadi integer dimana 1 untuk jawaban “yes” dan 0 untuk jawaban “no”
12	<b>Penyembuhan luka yang lebih lambat (<i>Delayed healing</i>)</b>	Apakah pasien mengalami penyembuhan luka yang lambat. Tipe data awal <i>object</i> , akan diubah setelahnya menjadi integer dimana 1 untuk jawaban “yes” dan 0 untuk jawaban “no”
13	<b>Melemahnya sebagian otot (<i>Partial paresis</i>)</b>	Apakah terjadi pelemahan otot pada pasien. Tipe data awal <i>object</i> , akan diubah setelahnya menjadi integer dimana 1 untuk jawaban “yes” dan 0 untuk jawaban “no”
14	<b>Kekakuan otot (<i>Muscle stiffness</i>)</b>	Apakah terjadi kekakuan otot pada pasien. Tipe data awal <i>object</i> , akan diubah setelahnya menjadi integer dimana 1 untuk jawaban “yes” dan 0 untuk jawaban “no”
15	<b>Kerontokan rambut (<i>Alopecia</i>)</b>	Apakah pasien mengalami kerontokan rambut. Tipe data awal <i>object</i> , akan diubah setelahnya menjadi integer dimana 1 untuk jawaban “yes” dan 0 untuk jawaban “no”
16	<b>Obesitas (<i>Obesity</i>)</b>	Apakah pasien termasuk obesitas. Tipe data awal <i>object</i> , akan diubah setelahnya menjadi integer dimana 1 untuk jawaban “yes” dan 0 untuk jawaban “no”
17	<b>Kelas klasifikasi (<i>Class</i>)</b>	Hasil klasifikasi pasien. Tipe data awal <i>object</i> , akan diubah setelahnya menjadi integer dimana 1 untuk hasil “ <i>positive</i> ” dan 0 untuk hasil “ <i>negative</i> ”

Contoh 5 data terakhir yang diambil dari *dataset* diabetes tahap awal yang mendeskripsikan baris dari tiap atribut.

Tabel 3.2 Tabel dataset diabetes tahap awal atribut 1-8

<i>Age</i>	<i>Gender</i>	<i>Polyuria</i>	<i>Polydipsia</i>	<i>sudden weight loss</i>	<i>weakness</i>	<i>Polyphagia</i>	<i>Genital thrush</i>
39	<i>Female</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>	<i>No</i>

48	<i>Female</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>No</i>
58	<i>Female</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>No</i>
32	<i>Female</i>	<i>No</i>	<i>No</i>	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>No</i>
42	<i>Male</i>	<i>No</i>	<i>No</i>	<i>No</i>	<i>No</i>	<i>No</i>	<i>No</i>

Tabel 3.3 Tabel *dataset* diabetes tahap awal atribut 9-17

<i>visual blurring</i>	<i>Itching</i>	<i>Irritability</i>	<i>delayed healing</i>	<i>partial paresis</i>	<i>muscle stiffness</i>	<i>Alopecia</i>	<i>Obesity</i>	<i>Class</i>
<i>No</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>	<i>Yes</i>	<i>No</i>	<i>No</i>	<i>No</i>	<i>Positive</i>
<i>No</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>No</i>	<i>No</i>	<i>No</i>	<i>Positive</i>
<i>Yes</i>	<i>No</i>	<i>No</i>	<i>No</i>	<i>Yes</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>	<i>Positive</i>
<i>Yes</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>Negative</i>
<i>No</i>	<i>No</i>	<i>No</i>	<i>No</i>	<i>No</i>	<i>No</i>	<i>No</i>	<i>No</i>	<i>Negative</i>

Tabel 3.4 Tabel Distribusi Dataset Kategori *Class*

<i>Class</i>	<i>Count</i>
<b>Negative</b>	38.46%
<b>Positive</b>	61.54%

Tabel 3.4 tersebut menunjukkan distribusi persentase dari kategori 'class'. Ini adalah representasi dari proporsi kategori 'Negative' dan 'Positive' di seluruh dataset. Dari seluruh data, 38,46% tergolong dalam kategori 'Negative', yang berarti mereka tidak memiliki diabetes. Sementara itu, 61,54% dari data tergolong dalam kategori 'Positive', yang berarti mereka didiagnosis memiliki diabetes.

Tabel 3.5 Tabel Distribusi Dataset Kategori *Gender* Terhadap *Class*

<i>Gender</i>	<i>Class</i>	
	<i>Negative</i>	<i>Positive</i>
Female	9.89%	90.10%
Male	55.18%	44.81%

Tabel 3.5 menggambarkan distribusi persentase kasus diabetes berdasarkan gender, yaitu 'Female' (wanita) dan 'Male' (pria). Ini menunjukkan seberapa sering

kategori 'Negative' dan 'Positive' muncul dalam masing-masing kategori gender. Dari seluruh wanita yang ada dalam dataset, 9,90% didiagnosis tidak memiliki diabetes (kategori 'Negative') sementara 90,10% lainnya didiagnosis memiliki diabetes (kategori 'Positive'). Untuk pria dalam dataset, 55,18% didiagnosis tidak memiliki diabetes dan 44,82% didiagnosis memiliki diabetes. Dengan demikian, berdasarkan data yang ada, proporsi wanita yang didiagnosis memiliki diabetes jauh lebih tinggi dibandingkan dengan pria.

### ***3.6 Pre-Processing***

Tahap *Pre-processing* atau tahap pra-pemrosesan adalah langkah penting dalam mengkaji data yang akan digunakan, dengan tujuan untuk mengoptimalkan hasil yang diperoleh. Dalam penelitian ini, peneliti menerapkan tiga langkah utama dalam pra-pemrosesan untuk memastikan kualitas data yang akan diolah lebih lanjut. Berikut adalah langkah-langkah yang dijalankan dalam tahap pra-pemrosesan:

#### **3.6.1 Mengubah Tipe Data Non-Numerik menjadi Numerik (*Encoding*)**

Proses ini merupakan bagian dari pre-processing yang bertujuan untuk mengubah tipe data pada fitur dataset yang sebelumnya memiliki tipe data non-numerik, yaitu object, menjadi tipe data numerik (integer). Tujuan dari perubahan ini adalah untuk memudahkan dan menyederhanakan proses pengolahan data pada tahap berikutnya, sehingga memungkinkan proses analisis yang lebih efisien dan akurat dalam penelitian yang sedang dilakukan. Pada dataset diabetes tahap awal, terdapat 16 atribut atau fitur selain atribut usia yang memiliki tipe data non-numerik

(tipe data object). Salah satu contohnya pada atribut kelamin yang memiliki nilai female atau male, juga pada atribut Polyuria yang memiliki nilai yes atau no. Berikut adalah perubahan nilai tipe data non-numerik menjadi tipe data numerik.

Tabel 3.6 Hasil *Encoding*, Ubah Tipe Data Non-Numerik menjadi Numerik

<b>Data awal atribut jenis kelamin (Sex)</b>	<b>Data ubah atribut jenis kelamin (Sex)</b>
Pria	1
Wanita	0

Pada atribut jenis kelamin (*sex*), nilai pria diubah menjadi numerik dengan nilai 1 dan nilai wanita diubah menjadi numerik dengan nilai 0.

### 3.6.2 Penskalaan Data

Penskalaan data merupakan langkah krusial dalam pra-pemrosesan data untuk penelitian ini. Fitur 'Age' memiliki rentang perbedaan nilai yang signifikan dibandingkan dengan fitur lainnya, hal ini dapat mendominasi proses pembelajaran, yang mengarah pada hasil pelatihan yang kurang optimal. Dilakukannya penskalaan dapat mengatasi masalah ini.

Neural network umumnya sensitif terhadap fitur masukannya. Adanya perbedaan skala fitur dapat memperlambat proses pembelajaran karena metode gradient descent (yang digunakan dalam backpropagation) memerlukan waktu lebih lama untuk konvergensi ketika berhadapan dengan fitur yang memiliki skala berbeda-beda. Pada optimasi dengan ACO, memiliki fitur dengan skala yang serupa dapat membantu dalam mencapai optimasi yang lebih seimbang. Hal ini dikarenakan fitur dengan rentang yang lebih besar dapat memberikan pengaruh yang tidak proporsional terhadap proses optimasi.

Pada pra-pemrosesan penelitian ini, digunakan MinMaxScaler, sebuah pustaka Python yang berfungsi untuk menormalisasi fitur numerik. MinMaxScaler mengubah skala data sehingga nilai setiap fitur berada dalam rentang antara nol dan satu. Dalam konteks dataset diabetes yang digunakan, kolom 'Age' sebagai fitur numerik memiliki variasi nilai yang luas. Penerapan MinMaxScaler pada kolom 'Age' memastikan bahwa nilai usia ini tidak memberikan pengaruh yang berlebih saat melatih model machine learning, khususnya neural network. Tujuan utama dari normalisasi ini adalah untuk menghindari dominasi salah satu fitur numerik, seperti 'Age', dalam proses pembelajaran karena rentang nilainya yang lebih luas dibanding fitur lain. Rumus matematika dari perhitungan tradisional MinMaxScaler:

$$Skala = \frac{(X - X_{min})}{(X_{max} - X_{min})} \quad (3.1)$$

$X$  = Nilai asli yang ingin diskalakan

$X_{min}$  = Nilai minimum dalam seluruh dataset

$X_{max}$  = Nilai maksimum dalam seluruh dataset

Contoh Perhitungan:

Misalnya terdapat data usia / Age = 40, 58, 41, 45, 23, 31.  $X_{min} = 23$ ,

$X_{max} = 58$ . Maka:

$$Skala_{40} = \frac{(40 - 23)}{(58 - 23)} = \frac{17}{35} = 0.486$$

$$Skala_{58} = \frac{(58 - 23)}{(58 - 23)} = \frac{35}{35} = 0.486$$

$$Skala_{41} = \frac{(41 - 23)}{(58 - 23)} = \frac{18}{35} = 0.514$$

$$Skala_{45} = \frac{(45 - 23)}{(58 - 23)} = \frac{22}{35} = 0.629$$

$$Skala_{23} = \frac{(23 - 23)}{(58 - 23)} = \frac{0}{35} = 0.000$$

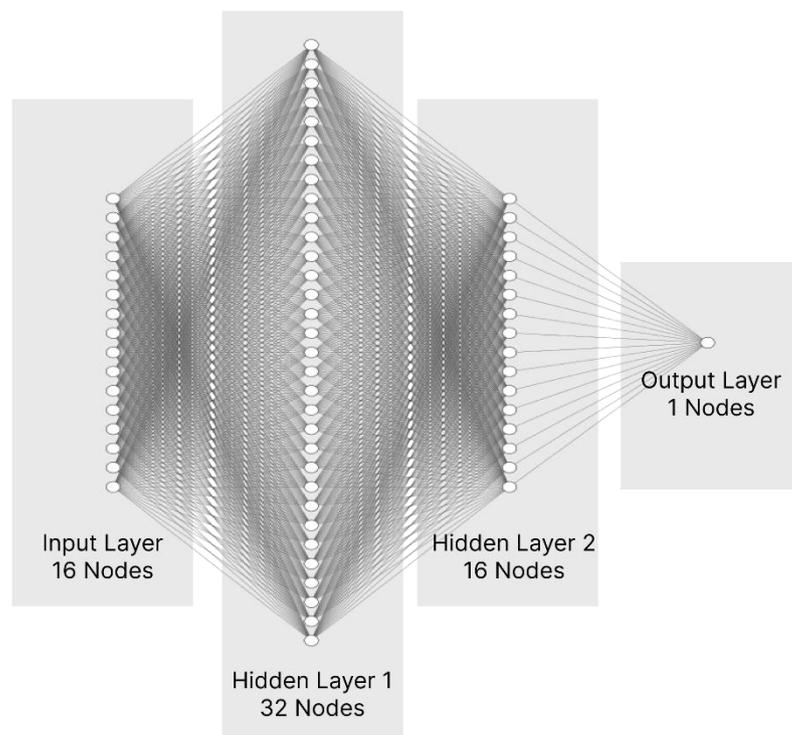
$$Skala_{31} = \frac{(31 - 23)}{(58 - 23)} = \frac{8}{35} = 0.229$$

### 3.6.3 Split Data

Pra-pemrosesan berikutnya adalah membagi data menjadi data fitur (X) dan label (y), kemudian membaginya lagi menjadi data latih dan data uji untuk melakukan pelatihan dan evaluasi model. Dampak dari data yang dibagi untuk kinerja model sangat signifikan karena akurasi suatu model bergantung dari kualitas data yang digunakan untuk pelatihan dan pengujian tersebut. Pembagian yang dilakukan ini untuk memastikan bahwa model tidak *overfitting* atau *underfitting*. Selain itu, membagi data juga membantu dalam mengevaluasi kinerja model, sehingga memungkinkan kita untuk membuat penyesuaian yang diperlukan guna meningkatkan akurasi. Pada *preprocess dataset* diabetes tahap awal, data dibagi menjadi data latih dan data uji dengan rasio 80% untuk data latih dan 20% untuk data uji.

### 3.7 Artificial Neural Network

*Artificial neural network* adalah model komputasi yang terinspirasi dari cara kerja sistem saraf biologis yaitu otak manusia. ANN terdiri dari unit-unit pemrosesan yang disebut *neuron (node)*, yang dihubungkan melalui konektor bernama sinapsis. Tujuan utama ANN adalah untuk meniru kemampuan belajar dan penalaran manusia dalam mengatasi berbagai tugas, seperti klasifikasi, regresi, pengenalan pola, dan sebagainya.



Gambar 3.4 Arsitektur ANN yang digunakan pada Penelitian

Berikut adalah penjelasan tentang cara menggunakan ANN, dan tahapannya:

1. Membangun jaringan: Dibangun jaringan syaraf dengan input layer, dua hidden layer dengan masing-masing layer memiliki jumlah nodesnya,

serta satu output layer. Pada *hidden layer* digunakan fungsi aktivasi ReLU (*Rectified Linear Unit*) dan pada *output layer* digunakan fungsi aktivasi Sigmoid. ReLU adalah fungsi aktivasi non-linear yang mengubah semua nilai input negatif menjadi nol dan mempertahankan semua nilai positif.

$$ReLU(x) = \max(0, x) \quad (3.2)$$

2. Sementara Sigmoid adalah fungsi aktivasi yang mengubah input menjadi nilai antara 0 dan 1. Karena outputnya terbatas, fungsi ini sering digunakan dalam klasifikasi biner, cocok untuk perhitungan output model pada dataset diabetes yang classnya hanya negatif atau positif.

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

3. Meminimalkan fungsi kerugian dengan kompilasi model menggunakan fungsi dari library TensorFlow yaitu optimizer 'Adam' dimana learning rate nya adalah 0.01. Fungsi 'binary\_crossentropy' digunakan juga karena cocok untuk klasifikasi biner. Fungsi loss (kerugian) menentukan seberapa jauh prediksi model dari nilai sebenarnya.

### ***3.8 Ant Colony Optimization***

*Ant Colony Optimization* (ACO) adalah teknik metaheuristik yang terinspirasi dari perilaku semut untuk mencari makanan. Dalam pencariannya,

semut menyebarkan feromon, zat kimia yang membantu semut lain dalam koloni untuk menemukan sumber makanan. Dalam konteks algoritma, feromon berfungsi sebagai penanda kualitas dari suatu jalur.

Untuk menerapkan algoritma ACO ke dalam model ANN untuk mengoptimalkan bobotnya, perlu dilakukan beberapa perubahan untuk menyesuaikan penggunaan fungsi ACO terhadap ANN. Perubahan ini masih menggunakan fungsi-fungsi dasar pada ACO. Beberapa hal yang perlu dimodifikasi:

1. Pada ACO untuk optimasi bobot, tidak ada panjang jarak yang dilalui semut dari titik  $i$  ke titik  $j$ , dikarenakan arsitektur neural network memiliki jarak yang tidak pasti dari node  $i$  ke node  $j$ . Oleh karenanya, ketika memilih probabilitas untuk semut memilih jalur  $i$  ke  $j$ , hanya intensitas feromon yang digunakan. Rumus Probabilitas Pemilihan Jalur  $i$  ke  $j$  menjadi:

$$P_{ij} = \frac{\tau_{ij}^{\alpha}}{\sum \tau_{ij}^{\alpha}} \quad (3.4)$$

2. Rumus penguapan feromon masih sama dengan mengetahui variabel  $\rho$  sebagai laju penguapan feromon dan  $\Delta\tau_{ij}$  sebagai jumlah total feromon yang ditambahkan ke jalur oleh semut. Untuk  $\Delta\tau_{ij}$ , konstanta  $Q$  nya adalah jumlah feromon yang ditinggalkan oleh semut pada jalur, sementara  $L$  nya diganti sebagai error.
3. Semut menambahkan feromon ke jalur yang mereka tempuh berdasarkan dari kualitas solusi. Feromon ditambahkan dengan rumus :

$$\Delta\tau_{ij} = \frac{Q}{Error} \quad (3.5)$$

$\Delta\tau_{ij}$  = Perubahan feromon yang ditambahkan oleh semut ke jalur  $i$  ke  $j$

$Q$  = Konstanta jumlah feromon yang ditinggalkan semut

Error = Cost error dari solusi yang ditemukan oleh semut

4. Pembaruan global feromon dari solusi terbaik tidak diimplementasikan secara eksplisit. Pembaruan feromon masih menggunakan rumus yang sama dengan rumus pembaruan feromon:

$$\tau_{ij} = (1 - \rho) \times \tau_{ij} + \Delta\tau_{ij} \quad (3.6)$$

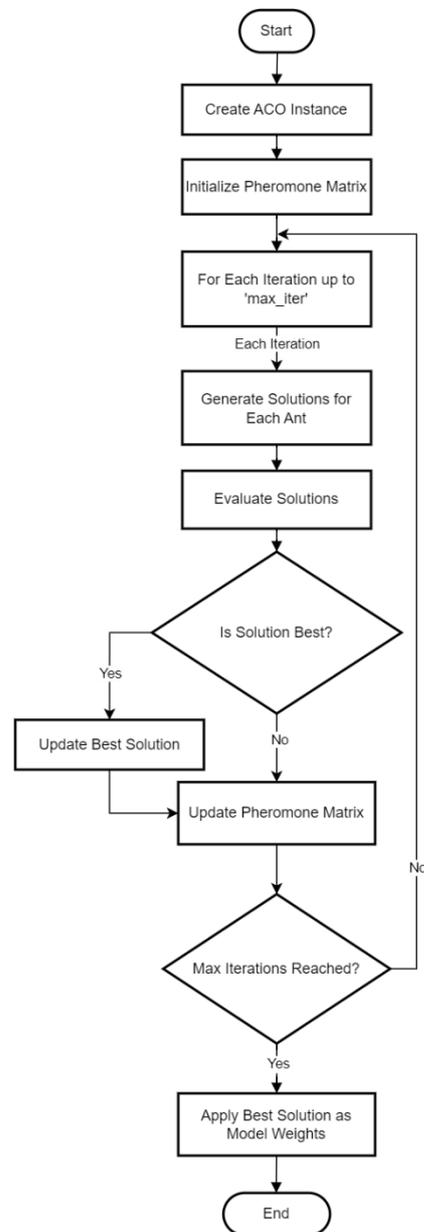
Parameter-parameter seperti jumlah semut, jumlah iterasi, alpha, beta, rho, dan Q ditentukan ketika akan menggunakan model ACO dengan memanggil class ACO yang sudah dibangun. Karena panjang jarak yang dilalui semut untuk optimasi bobot ANN tidak diketahui dan tidak dihitung, maka beta tidak digunakan meski nilainya disimpan di dalam model ACO.

1. **'n\_ants'**: Parameter yang menentukan jumlah semut (virtual) dalam simulasi pembangunan model. Jumlah semut berpengaruh pada jumlah solusi yang dijelajahi dalam setiap iterasi. Lebih banyak semut berarti lebih banyak jalur yang dijelajahi, yang dapat meningkatkan kemungkinan menemukan solusi optimal, tetapi juga mengharuskan lebih banyak waktu dan sumber daya komputasi.
2. **'max\_iter'**: Parameter yang menetapkan jumlah maksimum iterasi yang akan dilakukan algoritma. Limit iterasi memengaruhi lamanya algoritma

berjalan. Iterasi yang lebih banyak memungkinkan eksplorasi solusi yang lebih mendalam tetapi memerlukan waktu komputasi yang lebih lama. Jika terlalu sedikit, algoritma mungkin tidak mencapai solusi terbaik.

3. **'alpha ( $\alpha$ )'**: Parameter yang mengontrol pengaruh informasi feromon pada keputusan jalur yang diambil oleh semut. Nilai alpha yang tinggi meningkatkan pengaruh jejak feromon, membuat semut lebih cenderung mengikuti jejak dengan konsentrasi feromon yang lebih tinggi. Sebaliknya, nilai yang rendah membuat keputusan lebih acak, meningkatkan eksplorasi tetapi mengurangi kecepatan konvergensi.
4. **'beta ( $\beta$ )'**: Parameter yang mengontrol seberapa kuat informasi heuristik memengaruhi keputusan semut. Dengan nilai beta yang tinggi, semut akan lebih cenderung mengambil jalur dengan nilai heuristik yang menguntungkan (misalnya, jarak yang lebih pendek di masalah TSP). Ini mempromosikan eksploitasi jalur "baik" yang sudah ditemukan tetapi bisa mengurangi eksplorasi jalur baru.
5. **'rho ( $\rho$ )'**: Parameter koefisien dari penguapan feromon, yang menentukan seberapa cepat feromon menguap pada jalur yang telah dilalui semut. Nilai rho yang tinggi mengakibatkan penguapan yang cepat pada feromon, yang mendorong eksplorasi solusi baru dengan mengurangi peluang semut terjebak dalam solusi lokal terbaik. Nilai rho yang rendah membuat feromon bertahan lebih lama, mendorong semut untuk mengikuti jalur yang sama tetapi berisiko kurang menemukan solusi alternatif yang mungkin lebih baik.

6. **'Q'**: Parameter konstanta yang digunakan untuk menghitung jumlah feromon yang ditinggalkan oleh semut pada jalur, biasanya terkait dengan kualitas solusi. Nilai Q yang lebih tinggi meningkatkan jumlah feromon yang ditinggalkan, yang memperkuat jalur bagi semut lain dalam iterasi berikutnya. Ini bisa mempercepat konvergensi ke solusi yang baik, tetapi terlalu tinggi bisa mengurangi eksplorasi dan semut bisa terjebak pada solusi suboptimal.



Gambar 3.5 Diagram *Flowchart* Optimasi *ACO*

Tahapan-tahapan pada gambar 3.5:

1. **Create ACO Instance:** Proses optimasi dimulai dari membuat instance dari *ACO class* dengan parameter-parameter tertentu seperti jumlah semut

( $n\_ants$ ), iterasi maksimal ( $max\_iter$ ), dan koefisien algoritma ACO ( $alpha, beta, rho, q$ ).

2. **Initialize Pheromone Matrix:** Matriks feromon diinisialisasi. Feromon digunakan untuk para semut berkomunikasi menemukan jalur terbaik dalam menemukan makanan. Pada kasus ini, dibangun peta feromon *virtual* yang dapat membantu semut menemukan bobot *neural network* yang paling baik. Matriks feromon ini diinisialisasi menjadi dimensi yang sesuai dengan jumlah bobot pada model ANN ( $n\_weights$ ).
3. **For Each Iteration up to max\_iter:** Untuk Setiap Iterasi hingga ' $max\_iter$ ' sebuah loop dijalankan yang akan berlangsung hingga jumlah iterasi maksimal yang telah ditentukan.
4. **Generate Solutions for Each Ant:** Di dalam loop, setiap semut menghasilkan sebuah solusi. Hal ini dilakukan dengan memilih komponen selanjutnya (*choose next component*) atau probabilitas pemilihan jalur untuk semut memilih jalur berdasarkan intensitas feromon terhadap parameter *alpha*-nya.
5. **Evaluate Solutions:** Setiap solusi yang dihasilkan para semut dievaluasi dengan mengaplikasikan solusi-solusi tersebut sebagai bobot-bobot di model *neural network* dan kemudian dilakukan evaluasi performa model menggunakan dataset ( $X\_train, y\_train$ ). *Error* atau *loss* dari hasil evaluasi kemudian dikumpulkan untuk setiap solusi
6. **Update Best Solution:** Dilakukan pemeriksaan untuk menentukan apakah solusi saat ini memiliki kesalahan yang lebih rendah daripada kesalahan

terbaik (paling rendah) yang telah tercatat sejauh ini. Jika iya, maka dilakukan pembaruan terhadap (*best\_solution*) dan (*best\_error*) nya.

7. **Update Pheromone Matrix:** Setelah semua solusi telah dievaluasi dan solusi terbaik diperbarui, selanjutnya dilakukan pembaruan terhadap matriks feromon. Pembaruan ini didasarkan pada solusi yang dihasilkan dan kesalahan yang sesuai, menyesuaikan tingkat feromon untuk mengarahkan iterasi berikutnya menuju solusi yang lebih baik. Pembaruan ini melibatkan pengurangan tingkat feromon di semua jalur dengan faktor tertentu ( $1 - \rho$ , dikenal sebagai penguapan) dan kemudian meningkatkan tingkat feromon pada jalur yang sesuai dengan solusi yang lebih baik (*delta\_pheromone*).
8. **Not Max Iterations:** Dilakukan pemeriksaan untuk menentukan apakah jumlah iterasi maksimal telah tercapai. Jika belum, proses kembali ke *loop* untuk menghasilkan solusi baru untuk setiap semut.
9. **Apply Best Solution as Model Weights:** Dilakukan pemeriksaan untuk menentukan apakah jumlah iterasi maksimal telah tercapai. Jika belum, proses kembali ke *loop* untuk menghasilkan solusi baru untuk setiap semut.
10. **End of Process:** Setelah jumlah iterasi maksimal tercapai, solusi terbaik yang ditemukan selama iterasi diterapkan pada model jaringan saraf sebagai bobotnya.

### 3.9 Backpropagation

*Backpropagation*: Algoritma yang digunakan untuk memperbarui bobot dan bias berdasarkan kesalahan (*error*) antara prediksi ANN yang dihasilkan dan target yang diharapkan.

Pada penelitian ini, digunakan fungsi '**fit**' dari *library* TensorFlow yang menangani *Backpropagation* secara internal. *Backpropagation* ini menggunakan turunan parsial dan aturan rantai dari kalkulus agar dapat menghitung *gradient* fungsi biaya (*cost function*) terhadap setiap bobot di jaringan syaraf.

Gradien ini kemudian digunakan dalam algoritma optimasi, untuk penelitian ini digunakan Adam untuk menyesuaikan bobot dalam upaya meminimalkan kesalahan.

Beberapa parameter dalam Backpropagation:

1. '**epochs**': Menentukan bahwa seluruh *dataset* akan melewati jaringan syaraf (*forward pass* dan *backward pass*) sebanyak perulangan sekian kali. Setiap iterasi dari keseluruhan dataset disebut sebuah epoch.
2. '**batch\_size**': Menentukan jumlah sampel data yang akan disebarikan melalui jaringan (*forward pass*) sebelum melakukan *backward pass* untuk menyesuaikan bobot. Ini adalah metode *stochastic gradient descent* (dalam kasus ini, *mini-batch*) karena mengambil sampel dari dataset.
3. '**validation\_data**': Digunakan untuk mengevaluasi performa model pada set data yang tidak terlihat selama pelatihan, membantu dalam mengidentifikasi masalah seperti *overfitting*.

### 3.10 Akurasi Hasil

Untuk mengetahui akurasi serta membandingkan hasil akurasi kedua metode, digunakan *Confusion Matrix*.

		Actual Label	
		Positive (1)	Negative (0)
Predicted Label	Positive (1)	True Positive (TP)	False Positive (FP)
	Negative (0)	False Negative (FN)	True Negative (TN)

Gambar 3.6 *Confusion Matrix*

*Accuracy*, dari semua kelas negatif dan positif, berapa banyak yang terprediksi benar =  $\frac{TP+TN}{TP+FP+FN+TN}$

*Precision*, dari semua kelas yang diprediksi positif, berapa banyak yang benar bernilai positif =  $\frac{TP}{TP+FP}$

*Recall*, dari semua kelas positif, berapa banyak yang terprediksi benar =  $\frac{TP}{TP+FN}$

F-Measure, melakukan pengukuran terhadap recall dan precision bersamaan =  $2 \times \frac{Precision \times Recall}{Precision+Recall}$

Evaluasi seperti akurasi, presisi, *recall*, dan *F1-score* akan digunakan untuk mengukur seberapa baik akurasi perhitungan ANN yang telah dioptimasi.

## BAB IV

### UJI COBA DAN PEMBAHASAN

#### 4.1 Skenario Uji Coba

Dalam penelitian ini, digunakan data dengan 520 *instances*. Dilakukan percobaan dengan menggunakan tiga jenis pembagian data untuk pelatihan (*training*) dan validasi (*testing*). Model A menggunakan 60% data untuk pelatihan dan 40% untuk validasi. Model B menggunakan 70% data untuk pelatihan dan 30% untuk validasi. Model C menggunakan 80% data untuk pelatihan dan 20% untuk validasi.

Tabel 4.1 Perbandingan 4 Pembagian *Split* Data Latih dan Data Uji

Model	<i>Split</i>	<i>Train</i>	<i>Test</i>
A	60%:40%	312	208
B	70%:30%	364	156
C	80%:20%	416	104

Selanjutnya, setiap model tersebut akan dibandingkan lagi dengan empat jenis metode pelatihan berbeda. Metode pertama adalah *Neural Network* tanpa ACO maupun BP, dimana fungsi '*build\_nn*' digunakan untuk membuat model NN dengan tiga lapisan, termasuk dua lapisan tersembunyi (*hidden layers*). *Hidden layer* pertama memiliki 32 *neurons* dengan fungsi aktivasi '*relu*', *hidden layer* kedua memiliki 16 *neurons* dengan fungsi aktivasi yang sama, dan layer terakhir menghasilkan *output* dengan fungsi aktivasi '*sigmoid*'. Model ini dikompilasi

dengan *optimizer Adam* yang *learning rate*-nya adalah 0,01 dan untuk perhitungan *loss*-nya digunakan '*binary\_crossentropy*'.

Metode kedua adalah metode *Neural Network* dengan optimasi ACO (*Ant Colony Optimization*), dimana parameter yang digunakan adalah: jumlah semut sebanyak 5; iterasi maksimum 100; *alpha* (seberapa penting jejak feromon dalam mempengaruhi keputusan semut) 0,5; *beta* (mengontrol seberapa penting informasi heuristic (misalnya, jarak ke tujuan) dalam mempengaruhi keputusan semut) 2; *rho* (laju/koefisien penguapan feromon) 0,2, dan *q* (menghitung perubahan delta pheromone berdasarkan solusi dan error yang dihasilkan oleh semut) 0,7.

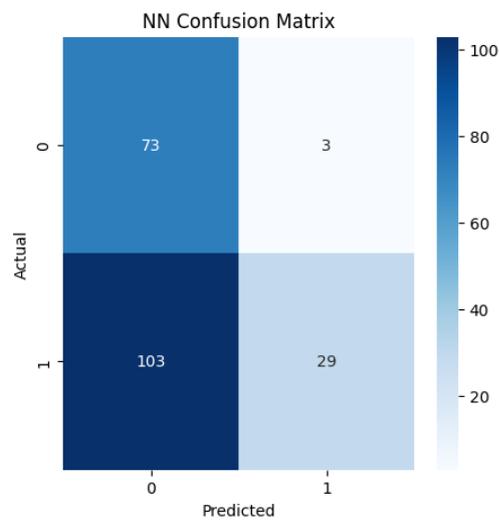
Metode ketiga, *Neural Network* dengan optimasi BP (*Backpropagation*) dimana pelatihan dilakukan yang dilakukan memiliki 100 *epochs* dan ukuran *batch* sebanyak 32, serta menggunakan data validasi. Metode keempat, yaitu *Neural Network* dengan optimasi ACO dan BP.

## 4.2 Hasil Uji Coba

### 4.2.1 MODEL A

Perbandingan data 60%:40%. Jumlah data training: 312; Jumlah data validasi: 208

#### 4.2.1.1 Hasil Uji Model A dengan *Neural Network*



Gambar 4.1 *Confusion Matrix* Model A, NN tanpa optimasi ACO dan BP

TP: 29, TN: 73, FP: 3, FN: 103

$$Accuracy = \frac{29 + 73}{29 + 73 + 3 + 103} = \frac{102}{208} = 0.4904 = 49.04\%$$

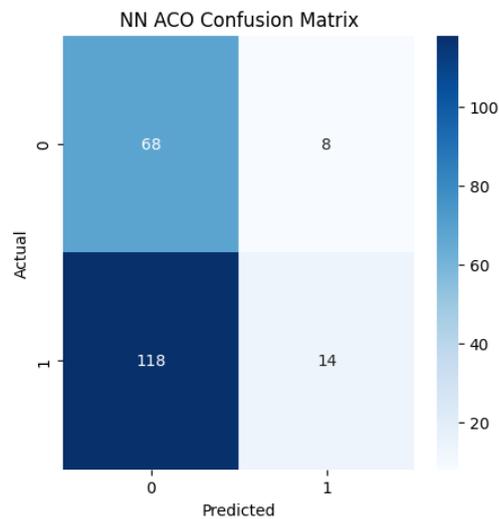
$$Precision = \frac{29}{29 + 3} = 0.9062 = 90.62\%$$

$$Recall = \frac{29}{29 + 103} = 0.2197 = 21.97\%$$

$$F1 = 2 \times \frac{0.9062 \times 0.2197}{0.9062 + 0.2197} = 0.3537 = 35.37\%$$

#### 4.2.1.2 Hasil Uji Model A dengan *Neural Network* dan *Ant Colony*

##### *Optimization*



Gambar 4.2 *Confusion Matrix* Model A, NN dioptimasi dengan ACO

TP: 14, TN: 68, FP: 8, FN: 118

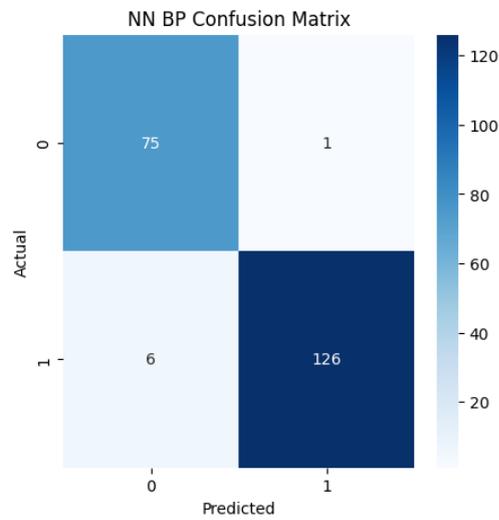
$$Accuracy = \frac{14 + 68}{14 + 68 + 8 + 118} = \frac{82}{208} = 0.3942 = 39.42\%$$

$$Precision = \frac{14}{14 + 8} = 0.6364 = 63.64\%$$

$$Recall = \frac{14}{14 + 118} = 0.1061 = 10.61\%$$

$$F1 = 2 \times \frac{0.6364 \times 0.1061}{0.6364 + 0.1061} = 0.1818 = 18.18\%$$

### 4.2.1.3 Hasil Uji Model A dengan *Neural Network* dan *Backpropagation*



Gambar 4.3 *Confusion Matrix* Model A, NN dioptimasi dengan BP

TP: 126, TN: 75, FP: 1, FN: 6

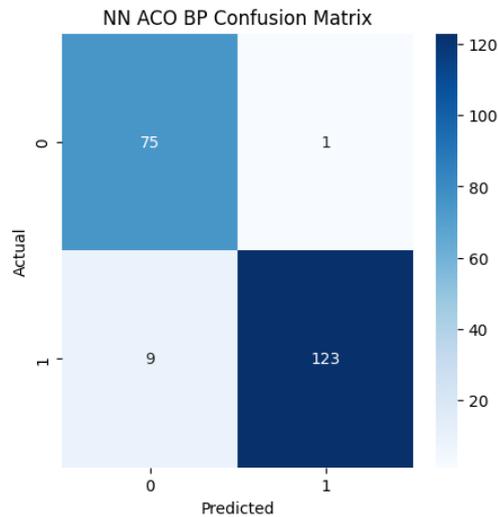
$$Accuracy = \frac{126 + 75}{126 + 75 + 1 + 6} = \frac{201}{208} = 0.9663 = 96.63\%$$

$$Precision = \frac{126}{126 + 1} = 0.9921 = 99.21\%$$

$$Recall = \frac{126}{126 + 6} = 0.9545 = 95.45\%$$

$$F1 = 2 \times \frac{0.9921 \times 0.9545}{0.9921 + 0.9545} = 0.9730 = 97.30\%$$

#### 4.2.1.4 Hasil Uji Model A dengan *Neural Network*, *Ant Colony Optimization*, dan *Backpropagation*



Gambar 4.4 *Confusion Matrix* Model A, NN dioptimasi dengan ACO dan BP

TP: 123, TN: 75, FP: 1, FN: 9

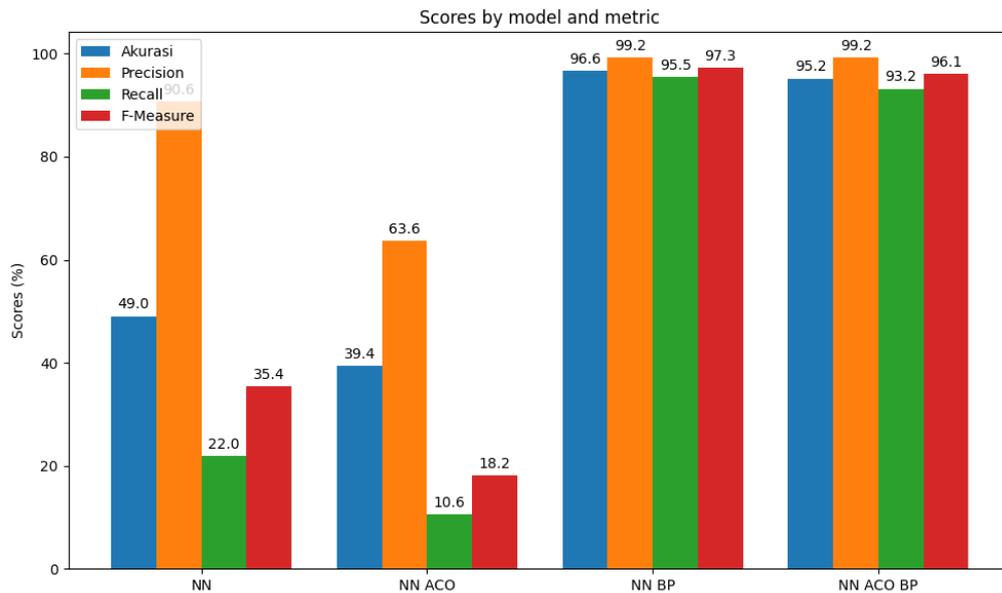
$$Accuracy = \frac{123 + 75}{123 + 75 + 1 + 9} = \frac{198}{208} = 0.9519 = 95.19\%$$

$$Precision = \frac{123}{123 + 1} = 0.9919 = 99.19\%$$

$$Recall = \frac{123}{123 + 9} = 0.9318 = 93.18\%$$

$$F1 = 2 \times \frac{0.9919 \times 0.9318}{0.9919 + 0.9318} = 0.9609 = 96.09\%$$

#### 4.2.1.5 Perbandingan *Confusion Matrix* dari Model A Terhadap 4 Pelatihan



Gambar 4.5 Perbandingan *Confusion Matrix* 4 Pelatihan pada Model A

Tabel 4.2 Perbandingan *Confusion Matrix* dari 4 Pelatihan pada Model A

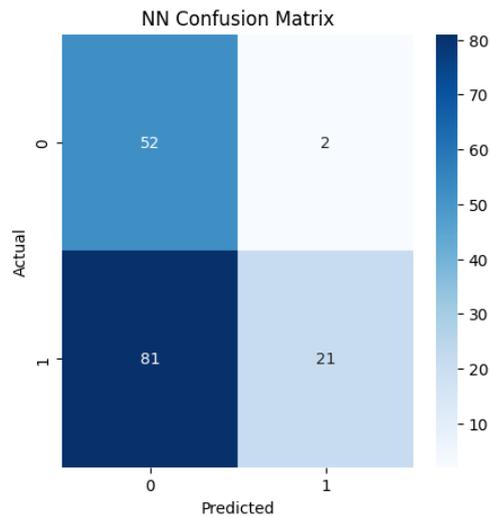
Model	<i>Performance (%)</i>			
	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>
NN	49.04	90.62	21.97	35.37
NN dioptimasi ACO	39.42	63.64	10.61	18.18
NN dioptimasi BP	96.63	99.21	95.45	97.30
NN dioptimasi ACO & BP	95.19	99.19	93.18	96.09

Dari Tabel 4.2, Akurasi tertinggi dari model A dihasilkan oleh *Neural Network* dengan *Backpropagation* dengan nilai mencapai 96.63%. Sementara itu, akurasi terendah dihasilkan oleh *Neural Network* dengan optimasi *Ant Colony Optimization* dengan nilai 39.42%.

## 4.2.2 MODEL B

Perbandingan data 70%:30%. Jumlah data training: 364; Jumlah data validasi: 156

### 4.2.2.1 Hasil Uji Model B dengan *Neural Network*



Gambar 4.6 *Confusion Matrix* Model B, NN tanpa optimasi ACO dan BP

TP: 21, TN: 52, FP: 2, FN: 81

$$Accuracy = \frac{21 + 52}{21 + 52 + 2 + 81} = \frac{73}{156} = 0.4679 = 46.79\%$$

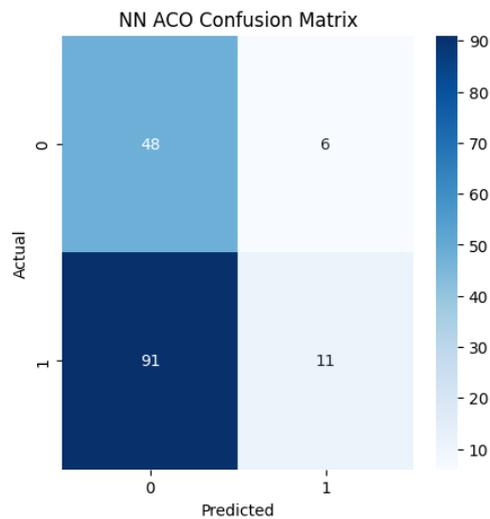
$$Precision = \frac{21}{21 + 2} = 0.9130 = 91.30\%$$

$$Recall = \frac{21}{21 + 81} = 0.2059 = 20.59\%$$

$$F1 = 2 \times \frac{0.9130 \times 0.2059}{0.9130 + 0.2059} = 0.3360 = 33.60\%$$

#### 4.2.2.2 Hasil Uji Model B dengan *Neural Network* dan *Ant Colony*

##### *Optimization*



Gambar 4.7 *Confusion Matrix* Model B, NN dioptimasi dengan ACO

TP: 11, TN: 48, FP: 6, FN: 91

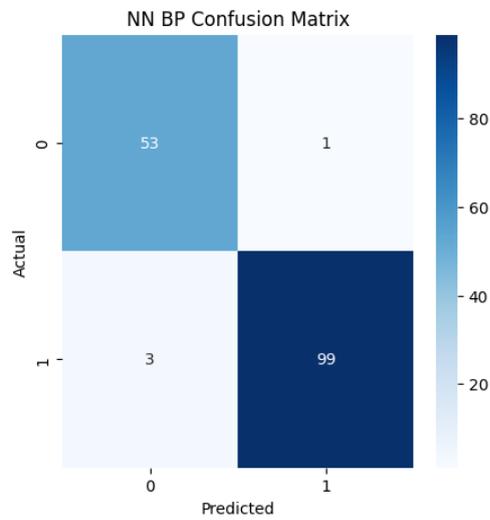
$$Accuracy = \frac{11 + 48}{11 + 48 + 6 + 91} = \frac{59}{156} = 0.3782 = 37.82\%$$

$$Precision = \frac{11}{11 + 6} = 0.6471 = 64.71\%$$

$$Recall = \frac{11}{11 + 91} = 0.1078 = 10.78\%$$

$$F1 = 2 \times \frac{0.6471 \times 0.1078}{0.6471 + 0.1078} = 0.1849 = 18.49\%$$

### 4.2.2.3 Hasil Uji Model B dengan *Neural Network* dan *Backpropagation*



Gambar 4.8 *Confusion Matrix* Model B, NN dioptimasi dengan BP

TP: 99, TN: 53, FP: 1, FN: 3

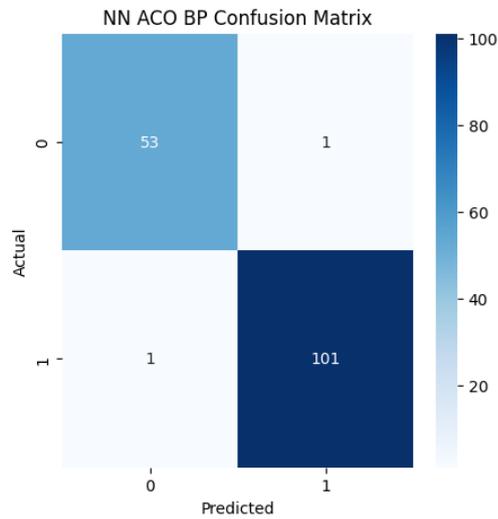
$$Accuracy = \frac{99 + 53}{99 + 53 + 1 + 3} = \frac{152}{156} = 0.9744 = 97.44\%$$

$$Precision = \frac{99}{99 + 1} = 0.9900 = 99.00\%$$

$$Recall = \frac{99}{99 + 3} = 0.9706 = 97.06\%$$

$$F1 = 2 \times \frac{0.9900 \times 0.9706}{0.9900 + 0.9706} = 0.9802 = 98.02\%$$

#### 4.2.2.4 Hasil Uji Model B dengan *Neural Network, Ant Colony Optimization, dan Backpropagation*



Gambar 4.9 *Confusion Matrix* Model B, NN dioptimasi dengan ACO dan BP

TP: 101, TN: 53, FP: 1, FN: 1

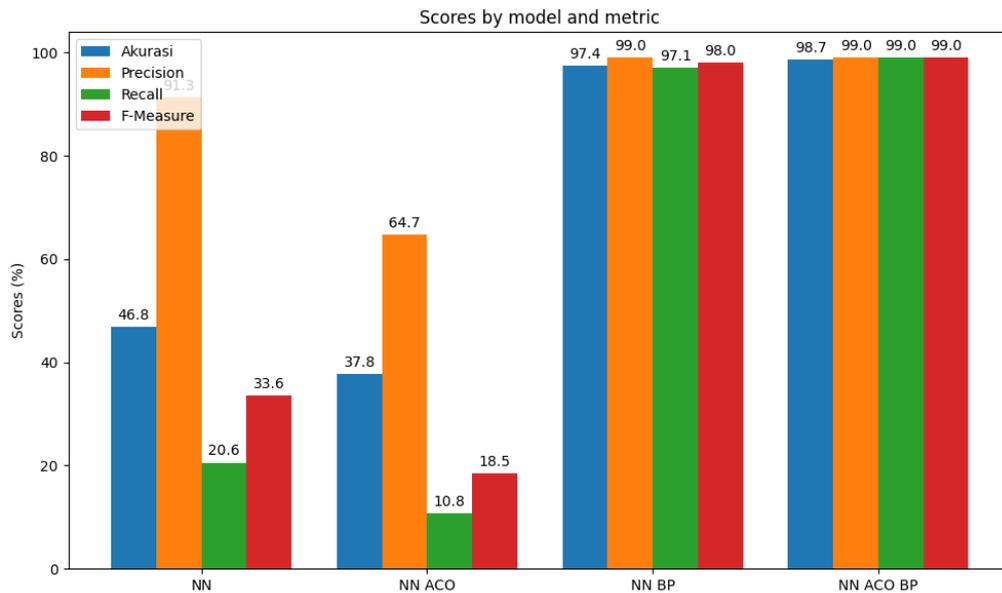
$$Accuracy = \frac{101 + 53}{101 + 53 + 1 + 1} = \frac{154}{156} = 0.9872 = 98.72\%$$

$$Precision = \frac{101}{101 + 1} = 0.9902 = 99.02\%$$

$$Recall = \frac{101}{101 + 1} = 0.9902 = 99.02\%$$

$$F1 = 2 \times \frac{0.9902 \times 0.9902}{0.9902 + 0.9902} = 0.9902 = 99.02\%$$

#### 4.2.2.5 Perbandingan *Confusion Matrix* dari Model B Terhadap 4 Pelatihan



Gambar 4.10 Perbandingan *Confusion Matrix* 4 Pelatihan pada Model B

Tabel 4.3 Perbandingan *Confusion Matrix* dari 4 Pelatihan pada Model B

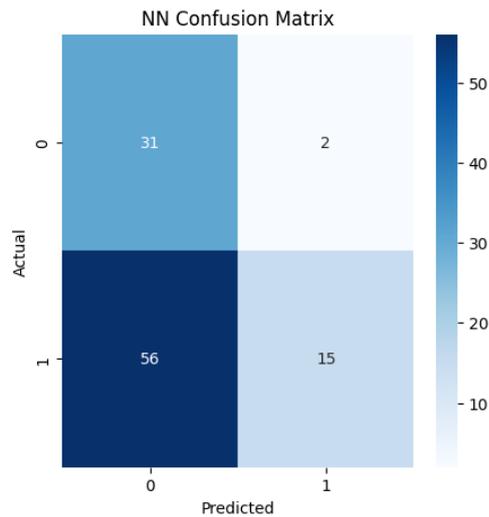
Model	<i>Performance (%)</i>			
	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>
NN	46.79	91.30	20.59	33.60
NN dioptimasi ACO	37.82	64.71	10.78	18.49
NN dioptimasi BP	97.44	99.00	97.06	98.02
NN dioptimasi ACO & BP	98.72	99.02	99.02	99.02

Berdasarkan tabel 4.3, Akurasi tertinggi dari model B dihasilkan oleh *Neural Network* dengan optimasi *Ant Colony Optimization* dan *Backpropagation* dengan nilai mencapai 98.72%. Sementara itu, akurasi terendah dihasilkan oleh *Neural Network* dengan optimasi *Ant Colony Optimization* dengan nilai 37.82%.

### 4.2.3 MODEL C

Perbandingan data 80%:20%. Jumlah data training: 416; Jumlah data validasi: 104

#### 4.2.3.1 Hasil Uji Model C dengan *Neural Network*



Gambar 4.11 *Confusion Matrix* Model C, NN tanpa optimasi ACO dan BP

TP: 15, TN: 31, FP: 2, FN: 56

$$Accuracy = \frac{15 + 31}{15 + 31 + 2 + 56} = \frac{46}{104} = 0.4423 = 44.23\%$$

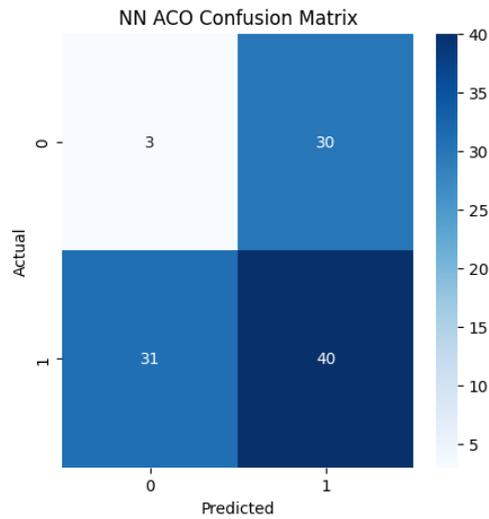
$$Precision = \frac{15}{15 + 2} = 0.8824 = 88.24\%$$

$$Recall = \frac{15}{15 + 56} = 0.2113 = 21.13\%$$

$$F1 = 2 \times \frac{0.8824 \times 0.2113}{0.8824 + 0.2113} = 0.3409 = 34.09\%$$

### 4.2.3.2 Hasil Uji Model C dengan *Neural Network* dan *Ant Colony Optimization*

#### *Optimization*



Gambar 4.12 *Confusion Matrix* Model C, NN dioptimasi dengan ACO

TP: 40, TN: 3, FP: 30, FN: 31

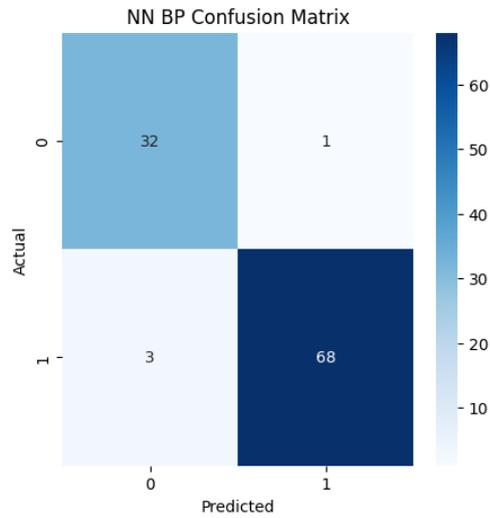
$$Accuracy = \frac{40 + 3}{40 + 3 + 30 + 31} = \frac{43}{104} = 0.4135 = 41.35\%$$

$$Precision = \frac{40}{40 + 30} = 0.5714 = 57.14\%$$

$$Recall = \frac{40}{40 + 31} = 0.5634 = 56.34\%$$

$$F1 = 2 \times \frac{0.5714 \times 0.5634}{0.5714 + 0.5634} = 0.5674 = 56.74\%$$

#### 4.2.3.3 Hasil Uji Model C dengan *Neural Network* dan *Backpropagation*



Gambar 4.13 *Confusion Matrix* Model C, NN dioptimasi dengan BP

TP: 68, TN: 32, FP: 1, FN: 3

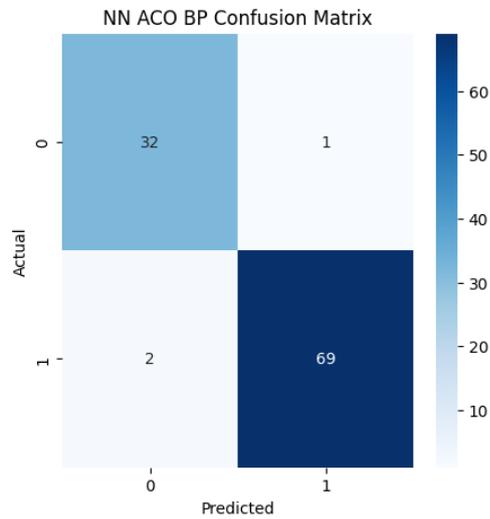
$$Accuracy = \frac{68 + 32}{68 + 32 + 1 + 3} = \frac{100}{104} = 0.9615 = 96.15\%$$

$$Precision = \frac{68}{68 + 1} = 0.9855 = 98.55\%$$

$$Recall = \frac{68}{68 + 3} = 0.9577 = 95.77\%$$

$$F1 = 2 \times \frac{0.9855 \times 0.9577}{0.9855 + 0.9577} = 0.9714 = 97.14\%$$

#### 4.2.3.4 Hasil Uji Model C dengan *Neural Network, Ant Colony Optimization, dan Backpropagation*



Gambar 4.14 *Confusion Matrix* Model C, NN dioptimasi dengan ACO dan BP

TP: 69, TN: 32, FP: 1, FN: 2

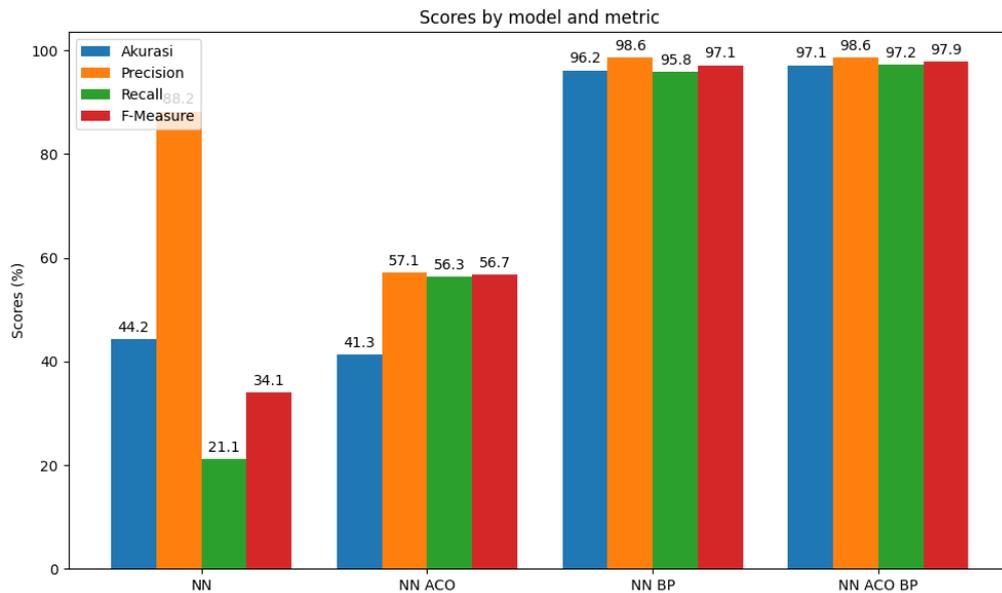
$$Accuracy = \frac{69 + 32}{69 + 32 + 1 + 2} = \frac{101}{104} = 0.9712 = 97.12\%$$

$$Precision = \frac{69}{69 + 1} = 0.9857 = 98.57\%$$

$$Recall = \frac{69}{69 + 2} = 0.9718 = 97.18\%$$

$$F1 = 2 \times \frac{0.9857 \times 0.9718}{0.9857 + 0.9718} = 0.9787 = 97.87\%$$

#### 4.2.3.5 Perbandingan *Confusion Matrix* dari Model C Terhadap 4 Pelatihan



Gambar 4.15 Perbandingan *Confusion Matrix* 4 Pelatihan pada Model C

Tabel 4.4 Perbandingan *Confusion Matrix* dari 4 Pelatihan pada Model C

Model	<i>Performance (%)</i>			
	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>
NN	44.23	88.24	21.13	34.09
NN dioptimasi ACO	41.35	57.14	56.34	56.74
NN dioptimasi BP	96.15	98.55	95.77	97.14
NN dioptimasi ACO & BP	97.12	98.57	97.18	97.87

Dari tabel 4.4, Akurasi tertinggi dari model C dihasilkan oleh *Neural Network* dengan optimasi *Ant Colony Optimization* dan *Backpropagation* dengan nilai mencapai 97.12%. Sementara itu, akurasi terendah dihasilkan oleh *Neural Network* dengan optimasi *Ant Colony Optimization* saja dengan nilai 41.35%.

### 4.3 Pembahasan

Pada penelitian "Penerapan Algoritma *Ant Colony Optimization* Dan *Backpropagation* Pada *Neural Network* Untuk Prediksi Diabetes Tahap Awal" ini, digunakan dataset yang diambil dari *UCI Machine Learning Repository* yaitu '*Early stage diabetes risk prediction dataset*' yang dipublikasikan pada 7 November 2020. Dataset tersebut memiliki 17 fitur dari 520 sampel pasien. Pada data tersebut terdapat 38.46% pasien negatif diabetes dan 61.54% pasien positif diabetes. Pada pasien wanita terdapat 9.89% pasien negatif diabetes dan 90.10% pasien positif diabetes. Adapun pada pasien pria, 55.18% negatif diabetes, dan 44.82% positif diabetes.

Digunakan model *Neural Network* dengan optimasi algoritma *Ant Colony Optimization* dan algoritma *Backpropagation* untuk melatih dataset tersebut. Pada uji coba, data dibagi menjadi 3 model pembagian *training* dan *testing* dengan pembagian *split data* 6:4, 7:3, dan 8:2. Masing-masing model kemudian dilatih dan dibandingkan dengan model yang terdiri dari *Neural Network* tanpa optimasi *Ant Colony Optimization* maupun *Backpropagation*, *Neural Network* dengan optimasi *Ant Colony*, *Neural Network* dengan optimasi *Backpropagation*, dan *Neural Network* dengan optimasi *Ant Colony Optimization* dan *Backpropagation*.

Hasil akurasi model terbaik dengan pelatihan model *Neural Network* yang dioptimasi *Ant Colony Optimization* dan *Backpropagation* ada pada model B. Model B menunjukkan kinerja terbaik di semua metrik, dengan akurasi 98.72%, presisi 99.02%, *recall* 99.02%, dan *F-measure* 99.02%, hal ini menunjukkan bahwa

alokasi 70% data untuk pelatihan dan 30% data untuk pengujian memberikan hasil terbaik untuk pelatihan dataset ini seperti yang ditunjukkan pada Tabel 4.5.

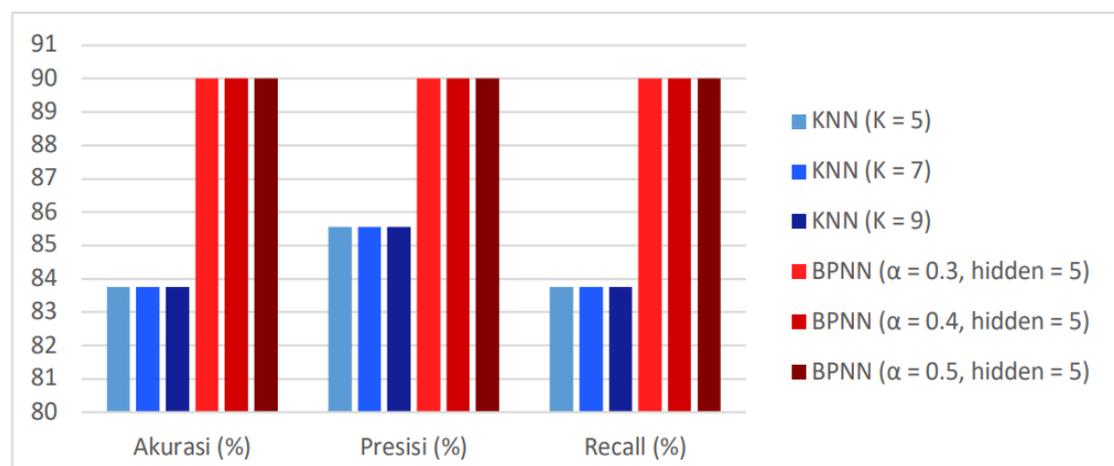
Tabel 4.5 Hasil Performa *Confusion Matrix* Seluruh Model

Model	Performance (%)					
	<i>Train</i>	<i>Test</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>
A	312	208	95.19	99.19	93.18	96.09
B	364	156	98.72	99.02	99.02	99.02
C	416	104	97.12	98.57	97.18	97.87

Hasil akurasi dari penelitian lain dengan dataset yang sama menunjukkan nilai yang bervariasi. Penelitian dengan judul "Perbandingan *K-Nearest Neighbor* dan *Backpropagation Neural Network* dalam Prediksi Resiko Diabetes Tahap Aawal" yang dilakukan oleh Jaka Permadi, Herfia Rhomadhona, dan Winda Aprianti dari Politeknik Negeri Tanah Laut. Penelitian ini berfokus pada perbandingan metode *K-Nearest Neighbor (KNN)* dan *Backpropagation Neural Network (BPNN)* dalam memprediksi resiko diabetes tahap awal.

Latar belakang penelitian menggarisbawahi pentingnya prediksi dini mengingat diabetes menjadi salah satu penyakit berbahaya di Indonesia, setelah stroke dan penyakit jantung koroner. Dataset yang digunakan berasal dari kuesioner pasien di rumah sakit *Sylhet, Bangladesh*. Hasil menunjukkan bahwa pada pengaturan  $K = 5$ ,  $K = 7$ , dan  $K = 9$ , pada KNN memiliki tingkat akurasi yang mencapai 83.75%, presisi 85.5497%, dan *recall* 83.75%. Sementara, BPNN dengan learning rate antara 0.3 hingga 0.5 dan jumlah node hidden sebanyak 5 unit mencapai tingkat akurasi, presisi, dan *recall* sebesar 90% seperti yang ada pada Gambar 4.16.

Kesimpulan dari penelitian ini adalah BPNN menunjukkan kemampuan yang lebih baik dibandingkan dengan KNN dalam memprediksi resiko diabetes tahap awal (Permadi et al., 2021).



Gambar 4.16 Hasil BPNN dan KNN penelitian Jaka Permadi dkk.

Penelitian dengan judul "Prediksi Kemungkinan Diabetes pada Tahap Awal Menggunakan Algoritma Klasifikasi *Random Forest*" dilakukan oleh Widya Apriliah, Ilham Kurniawan, Muhamad Baydhowi, dan Tri Haryati. Penelitian ini diterima pada 16 November 2020, direvisi pada 8 Desember 2020, dan diterima pada 14 Desember 2020. Dalam abstraknya, penelitian ini menyoroti bahwa penyakit diabetes merupakan penyakit kronis dengan pertumbuhan tercepat di dunia, dengan 422 juta penderita berdasarkan laporan dari WHO tahun 2018. Penelitian ini dilakukan dengan tujuan untuk menciptakan model yang dapat memprediksi kemungkinan terjadinya diabetes pada pasien dengan akurasi yang maksimal. Metode yang digunakan dalam penelitian ini meliputi *Support Vector Machine*, *Naive Bayes*, dan *Random Forest*. Data yang digunakan berasal dari *UCI Machine Learning Dataset* dari *Diabetes Hospital di Sylhet*, Bangladesh. Hasil dari

penelitian menunjukkan bahwa algoritma *Random Forest* memberikan performa terbaik dengan akurasi sebesar 97,88% dibandingkan dengan algoritma klasifikasi lainnya pada Tabel 4.9. Dalam kesimpulannya, penelitian ini menegaskan bahwa salah satu tantangan medis terpenting adalah mendeteksi diabetes pada tahap awal. Oleh karena itu, metode klasifikasi, khususnya algoritma *Random Forest*, dapat menjadi alat yang berharga untuk memprediksi atau mendiagnosis penyakit ini, serta meningkatkan otomatisasi analisis diabetes dibandingkan dengan algoritma *machine learning* lainnya (Aprilia et al., 2021).

Tabel 4.6 Perbandingan Algoritma Klasifikasi pada Penelitian Apriliah

<b>Classification Algorithm</b>	<b>Precision</b>	<b>Recall</b>	<b>F-Measure</b>	<b>Accuracy %</b>	<b>ROC</b>
SVM	0,949	0,948	0,948	94,80	0,949
Naive Bayes	0,925	0,919	0,920	91,92	0,964
Random Forest	<b>0,979</b>	<b>0,979</b>	<b>0,979</b>	<b>97,88</b>	<b>0,998</b>

Penelitian yang dilakukan oleh Özge Nur Ergün dan kawan-kawan dengan judul "Early Stage Diabetes Prediction Using Machine Learning Methods " berfokus pada penggunaan teknik machine learning untuk diagnosis diabetes tahap awal. Deteksi dini diabetes sangat penting untuk meningkatkan kualitas hidup pasien dan mengurangi komplikasi jangka panjang.

Studi ini menggunakan dataset diabetes umum dengan 16 fitur dari 520 individu untuk membangun model prediktif penelitian ini. Delapan metode *machine learning*, termasuk *Decision Tree*, *Random Forest*, *Support Vector Machine*, *XGBoost*, *K-nearest neighbor*, *Naive Bayes*, *Artificial Neural Network*, dan *Convolutional Neural Network*, digunakan pada dataset ini. Kinerja setiap

model divalidasi menggunakan skema 10-fold cross-validation, serta pengukuran *confusion matrix*.

*Convolutional Neural Network (CNN)*, khususnya varian satu dimensi, muncul sebagai model yang paling menonjol hasil kinerjanya, mencapai akurasi sebesar 99.04%, presisi 100%, recall 98.63%, dan skor F1 99.31%. Hasil ini menunjukkan bahwa model CNN 1D yang dikembangkan dalam penelitian ini bisa digunakan secara efektif untuk mengidentifikasi pasien diabetes pada tahap awal hanya dengan menanyakan serangkaian pertanyaan kepada mereka.

Penelitian ini juga mengeksplorasi pentingnya fitur-fitur terhadap klasifikasi diabetes, mengidentifikasi usia dan polidipsia (rasa haus yang berlebihan) sebagai parameter paling penting untuk mendeteksi penyakit.

Tabel 3.7 Delapan *Machine Learning* Penelitian Özge Nur Ergün dkk.

Method	Accuracy(%)	Precision(%)	Recall(%)	F1 Score(%)
K-Nearest Neighbor	89.42	88.92	90.34	89.09
Decision Tree	95.58	95.17	95.58	95.34
Random Forest	97.69	97.69	97.53	97.57
Support Vector Machine	94.62	94.18	94.51	94.31
Naïve Bayes	88.85	88.30	88.11	88.16
XGBoost	97.89	97.80	97.77	97.76
Artificial Neural Network	92.31	91.82	92.40	92.07
Convolutional Neural Network	<b>99.04</b>	<b>100.00</b>	<b>98.63</b>	<b>99.31</b>

Tabel 4.7 Perbandingan dengan Peneliti Lain dari Dataset yang Sama

Penelitian	Judul	Metode	Akurasi	Presisi	Recall
Jaka Permadi, dkk.	Perbandingan K-Nearest Neighbor Dan Backpropagation Neural Network Dalam Prediksi Resiko Diabetes Tahap Awal	Backpropagation Neural Network (BPNN)	90.00%	90.00%	90.00%
Widya Apriliah, dkk.	Prediksi Kemungkinan Diabetes Pada Tahap Awal Menggunakan Algoritma Klasifikasi Random Forest	Random Forest (RF)	97.88%	97.90%	97.90%
Özge Nur Ergün dkk.	Early Stage Diabetes Prediction Using Machine Learning Methods	Convolutional Neural Network	99.04%	100.00%	98.63%
<b>Peneliti</b>	<b>Penerapan Algoritma Ant Colony Optimization Dan Backpropagation Pada Neural Network Untuk Prediksi Diabetes Tahap Awal</b>	<b>Neural Network optimized with Ant Colony Optimization and Backpropagation</b>	<b>98.72%</b>	<b>99.02%</b>	<b>99.02%</b>

Jaka Permadi, dkk. mengambil pendekatan yang sedikit berbeda dengan membandingkan metode *K-Nearest Neighbor* dan *Backpropagation Neural Network* dalam penelitian mereka yang berjudul "Perbandingan *K-Nearest Neighbor* dan *Backpropagation Neural Network* Dalam Prediksi Resiko Diabetes Tahap Awal". Dari kedua metode tersebut, *Backpropagation Neural Network* (BPNN) terbukti memberikan hasil akurasi, presisi, dan recall sebesar 90.00%.

Widya Aprilia, dkk. dalam penelitiannya yang berjudul "Prediksi Kemungkinan Diabetes Pada Tahap Awal Menggunakan Algoritma Klasifikasi *Random Forest*", memanfaatkan metode *Random Forest* (RF) untuk prediksi. Dengan menggunakan algoritma ini, mereka berhasil mencapai tingkat akurasi sebesar 97.88%, preseisi dan recall sebesar 97.90%.

Pada penelitian yang dilakukan oleh Özge Nur Ergün dan kawan-kawan dengan judul "*Early Stage Diabetes Prediction Using Machine Learning Methods*" digunakan delapan *metode machine learning* untuk melatih dataset diabetes tahap awal. Salah satu model yang performanya paling menonjol adalah CNN (*Convolutional Neural Network*) satu dimensi dengan akurasi 99.04%, presisi 100%, dan recall 98.63%.

Pada penelitian "*Penerapan Neural Network Dengan Optimasi Ant Colony Optimization Dan Backpropagation Untuk Membangun Model Prediksi Diabetes Tahap Awal*" ini, metode yang digunakan adalah kombinasi dari *Neural Network* yang dioptimasi dengan *Ant Colony Optimization + Backpropagation* (ACO + BP). Hasil akurasi dari metode kombinasi ini mencapai 98.72%, presisi dan *recall* sebesar 99.02%

Dari tabel 4.7 di atas, dapat dilihat bahwa variasi penggunaan metode machine learning pada masing-masing penelitian dengan dataset yang sama menghasilkan nilai akurasi yang cukup tinggi meski berbeda-beda nilainya. Penelitian yang dilakukan oleh Özge Nur Ergün memiliki nilai akurasi dan presisi paling tinggi daripada penelitian yang lain. Pada penelitian "*Penerapan Algoritma Ant Colony Optimization Dan Backpropagation Pada Neural Network Untuk*

Prediksi Diabetes Tahap Awal” didapatkan nilai akurasi yang juga cukup tinggi yaitu 98.72%. Nilai presisi dan *recall*-nya sama yaitu mencapai 99.02%.

Secara keseluruhan, dapat disimpulkan bahwa meskipun semua metode menunjukkan potensi yang baik, tetapi CNN lebih menonjol sebagai metode yang paling efektif dalam hal akurasi dan presisi. Sementara kombinasi Ant Colony Optimization dengan Backpropagation menunjukkan kinerja yang hampir setara, perlu adanya penelitian lanjutan untuk mengoptimalkan model-model ini lebih menyeluruh. Penggunaan metode-metode ini dalam praktik klinis bisa sangat berdampak dalam meningkatkan deteksi dini diabetes, yang pada akhirnya dapat meningkatkan kualitas hidup pasien dan mengurangi beban penyakit ini secara global.

#### **4.4 Integrasi Penelitian dalam Islam**

Dalam mengkaji kompleksitas penyakit dan upaya menemukan solusi penyembuhan, kita sering kali mengabaikan nilai mendasar yang telah diajarkan sejak lama. Dari perspektif spiritual dan tradisi, seperti yang dikemukakan dalam Hadits dari Ibn Abbas, Rasulullah SAW mengingatkan kita tentang dua anugerah besar yang sering diabaikan: kesehatan dan waktu luang. Dalam konteks kesehatan, penting untuk mengakui bahwa meskipun kita harus berupaya mencari pengobatan, sebagaimana Nabi menyarankan untuk berobat, ada aspek-aspek tertentu dari kesehatan yang tidak bisa diatasi, seperti proses penuaan alami yang disebut sebagai "penyakit tua."

Menggabungkan pemahaman ini dengan pengetahuan modern, kita dapat melihat betapa pentingnya penelitian dalam pengembangan metode deteksi dini,

terutama untuk penyakit seperti diabetes yang telah menjadi pandemi global. Salah satu pendekatan inovatif dalam penelitian kesehatan masa kini adalah penerapan jaringan saraf tiruan, atau *Neural Network*, yang dioptimalkan dengan teknik seperti *Ant Colony Optimization* dan *Backpropagation*. Tujuannya adalah untuk membangun model prediksi yang akurat dan efisien untuk mendeteksi diabetes pada tahap awal.

Konsep ini secara khusus mencerminkan ayat Al-Qur'an, "Dan apabila aku sakit, Dialah yang menyembuhkanku" (*Asy-Syu'ara: 80*), menegaskan bahwa meskipun kesembuhan adalah anugerah dari Yang Maha Kuasa, sebagai manusia, kita diharapkan untuk mengambil langkah-langkah proaktif dalam mengelola kesehatan kita. Melalui integrasi antara pengobatan tradisional dan teknologi canggih, kita dapat menciptakan sistem yang tidak hanya prediktif tetapi juga preventif, memberi individu kesempatan yang lebih baik untuk mengambil alih kesehatan mereka sebelum kondisi tersebut berubah menjadi kritis.

#### 4.4.1 Menjaga Kesehatan

حَدَّثَنَا الْعَبَّاسُ بْنُ عَبْدِ الْعَظِيمِ الْعَنْبَرِيُّ، حَدَّثَنَا صَفْوَانُ بْنُ عَيْسَى، عَنْ عَبْدِ اللَّهِ بْنِ سَعِيدِ بْنِ أَبِي هَنْدٍ، عَنْ أَبِيهِ، قَالَ سَمِعْتُ ابْنَ عَبَّاسٍ، يَقُولُ قَالَ رَسُولُ اللَّهِ — صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ — " نِعْمَتَانِ مَغْبُونٌ فِيهِمَا كَثِيرٌ مِنَ النَّاسِ الصَّحَّةُ وَالْفَرَاغُ " .

*Dari Ibn Abbas RA, Rasulullah SAW bersabda, "Ada dua nikmat yang banyak orang tertipu (tidak menyadarinya), yaitu sehat dan waktu luang." (Hadits Riwayat Al-Bukhari No. 6412, dari Ibnu 'Abbas)*

Ibnu Baththol menekankan bahwa seseorang tidak dapat dikatakan benar-benar memiliki keleluasaan waktu kecuali jika ia juga menikmati kesehatan fisik. Bagi mereka yang diberkahi dengan dua anugerah ini, waktu luang dan kesehatan, maka ia menyarankan untuk bersikap semangat dan waspada agar tidak terjebak dalam ketidakbersyukuran kepada Allah SWT atas karunia-Nya tersebut. Menurutnya, rasa syukur yang sejati termanifestasi melalui kepatuhan penuh pada perintah-perintah-Nya dan menghindari segala yang dilarang oleh-Nya. Kegagalan dalam mengapresiasi karunia ini sebagaimana mestinya adalah bentuk penipuan bagi diri sendiri.

Ibnul Jauzi, dalam pandangannya, mengamati bahwa sering kali manusia diberi kesehatan namun keleluasaan waktunya terenggut oleh kesibukan dunia. Di lain waktu, seseorang mungkin memiliki kebebasan waktu tetapi dirundung oleh kondisi kesehatan yang buruk. Dia menggarisbawahi bahwa apabila kedua elemen ini, waktu luang dan kesehatan, hadir bersamaan, manusia cenderung dihadapkan pada rintangan berupa kemalasan dalam menjalankan ibadah dan ketaatan. Orang-orang dalam situasi seperti ini, menurutnya, adalah mereka yang telah terpedaya.

Selanjutnya, Ibnu Jauzi menyampaikan pesan penting yang seharusnya direnungkan oleh semua orang. Dia menggambarkan dunia sebagai ladang di mana manusia bekerja untuk memanen hasil di akhirat. Dunia, menurutnya, adalah pasar di mana kita menjual barang dagangan kita, dan profit sejati dari transaksi ini akan diperoleh di akhirat. Beliau menegaskan bahwa orang-orang yang menggunakan waktu luang dan kesehatan mereka untuk melakukan perbuatan yang baik serta melakukan ketaatan, maka akan mendapatkan kebahagiaan sejati. Sebaliknya, mereka yang menggunakan anugerah ini untuk melakukan perbuatan dosa, adalah mereka yang sesungguhnya tertipu. Karena setelah keleluasaan waktu, kesibukan akan datang menghampiri, dan demikian pula setelah kesehatan, kondisi sakit yang tidak diinginkan akan menyusul.

#### 4.4.2 Pentingnya Pengobatan

حَدَّثَنَا بَشْرُ بْنُ مُعَاذِ الْعَقَدِيِّ، قَالَ حَدَّثَنَا أَبُو عَوَانَةَ، عَنْ زِيَادِ بْنِ عِلَاقَةَ، عَنْ أُسَامَةَ بْنِ شَرِيكَ، قَالَ قَالَتِ الْأَعْرَابُ يَا رَسُولَ اللَّهِ أَلَا تَنْدَاوِي قَالَ " نَعَمْ يَا عِبَادَ اللَّهِ تَدَاوُوا فَإِنَّ اللَّهَ لَمْ يَضَعْ دَاءً إِلَّا وَضَعَ لَهُ شِفَاءً أَوْ قَالَ دَوَاءً إِلَّا دَاءً وَاحِدًا " . قَالُوا يَا رَسُولَ اللَّهِ وَمَا هُوَ قَالَ " الْهَرَمُ " . قَالَ أَبُو عَيْسَى وَفِي الْبَابِ عَنِ ابْنِ مَسْعُودٍ وَأَبِي هُرَيْرَةَ وَأَبِي خُرَّامَةَ عَنْ أَبِيهِ وَابْنِ عَبَّاسٍ . وَهَذَا حَدِيثٌ حَسَنٌ صَحِيحٌ .

*"Wahai Rosululloh, apakah kita berobat?, Nabi bersabda, "berobatlah, karena sesungguhnya Allah tidak menurunkan penyakit, kecuali pasti menurunkan obatnya, kecuali satu penyakit (yang tidak ada obatnya)" mereka bertanya, "apa itu" ? Nabi bersabda, "Penyakit tua." (HR. Tirmidzi 2038, dan disahihkan oleh al-Albani dalam Sunan Ibnu Majah 3436)*

Ibnu Qayyim menyatakan bahwa Allah telah menciptakan obat untuk setiap penyakit yang ada, meskipun pengetahuan mengenai pengobatan tersebut tidak selalu terbuka lebar untuk umat manusia. Beliau menekankan bahwa dalam hadits tersebut, terkandung sebuah arahan untuk berusaha menyembuhkan diri, suatu tindakan yang tidak bertentangan dengan prinsip tawakal. Melakukan pengobatan diibaratkannya sama seperti mengganti rasa lapar, dahaga, panas, dan dingin dengan keadaan yang berlawanan, sebagai bagian dari usaha untuk mencapai keseimbangan. Ibnu Qayyim juga berpendapat bahwa pengertian tauhid yang sejati mencapai kesempurnaan melalui pemahaman dan penerapan hukum sebab-akibat yang telah ditetapkan oleh Allah, baik melalui ajaran syariat maupun takdir. Beliau menginterpretasikan pernyataan Nabi tentang keberadaan obat untuk setiap penyakit sebagai motivasi bagi pasien dan praktisi medis untuk tidak berhenti mencari dan meneliti pengobatan. Keyakinan bahwa ada obat yang mampu menyembuhkan penyakit dapat menciptakan semangat harapan dalam diri orang sakit, menggantikan keputusan dengan optimisme dan ketenangan.

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Penelitian ini telah berhasil mengembangkan model latih Jaringan Saraf Tiruan (*Artificial Neural Network*) untuk memprediksi dataset diabetes tahap awal dengan mengadopsi alur kerja yang sistematis. Proses penelitian ini dimulai dari mengidentifikasi masalah, diikuti oleh studi literatur, eksplorasi data, pra-pemrosesan data, dan implementasi teknik-teknik pelatihan model yang mencakup *Neural Network*, *Ant Colony Optimization* (ACO), dan *Backpropagation* (BP).

Dalam konteks penggunaan data, peneliti menggunakan dataset '*Early stage diabetes risk prediction*' dari *UCI Machine Learning Repository*, yang terdiri dari 520 sampel data individu dengan 17 fitur. Distribusi data menunjukkan prevalensi yang lebih tinggi terhadap pasien positif diabetes, terutama di kalangan perempuan.

Penelitian ini menguji tiga model pembagian data, yaitu 6:4, 7:3, dan 8:2 untuk pelatihan dan validasi. Berikutnya, model-model tersebut dievaluasi dengan empat metode pelatihan yang berbeda: *Neural Network* tanpa optimasi dari ACO dan BP, *Neural Network* dengan optimasi ACO, *Neural Network* dengan optimasi BP, dan *Neural Network* dengan optimasi ACO dan BP.

Hasil penelitian menunjukkan bahwa model B, dengan pembagian data 70:30 untuk pelatihan dan validasi, memberikan performa terbaik. Model ini menggunakan *Neural Network* yang dioptimasi dengan ACO dan BP, dapat mencapai akurasi sebesar 98.72%, presisi 99.02%, *recall* 99.02%, dan *F-measure*

99.02%. Hasil tersebut menunjukkan bahwa alokasi 70% data untuk pelatihan, 30% data untuk pengujian, dan penggunaan metode optimasi ACO dan BP pada pelatihan *Neural Network*, merupakan pendekatan yang paling efektif untuk dataset ini.

## 5.2 Saran

Penelitian yang telah dilakukan menawarkan hasil yang signifikan, namun peneliti sadar bahwa masih banyak kekurangan yang perlu diperbaiki dan terdapat berbagai macam pendekatan yang dapat dieksplorasi untuk meningkatkan hasil penelitian di masa mendatang. Berikut ini adalah beberapa saran yang dapat dipertimbangkan untuk penelitian selanjutnya:

- a. Peneliti berikutnya dapat mencoba menggunakan metode penskalaan lain seperti *StandardScaler* untuk menormalisasi fitur usia pada dataset. Penerapan *StandardScaler* dapat membantu dalam mengurangi bias yang disebabkan oleh perbedaan skala antar fitur dan dapat meningkatkan efektivitas model.
- b. *Tuning hyperparameter* juga langkah yang baik untuk mengoptimalkan performa model. Dalam konteks *Ant Colony Optimization* dan *Backpropagation*, eksplorasi lebih lanjut terhadap *hyperparameter* seperti jumlah semut, laju penguapan, atau tingkat pembelajaran dapat memberikan wawasan tentang bagaimana setiap parameter mempengaruhi performa model.
- c. Dianjurkan juga untuk mencoba arsitektur *Neural Network* yang berbeda. Mengingat bahwa arsitektur yang berbeda mungkin lebih cocok untuk jenis

data yang berbeda atau masalah khusus, eksplorasi tersebut dapat menemukan konfigurasi yang lebih optimal untuk prediksi diabetes.

- d. Melakukan *feature selection* untuk mengidentifikasi fitur-fitur yang memiliki korelasi paling tinggi dengan hasil yang diinginkan agar dapat meningkatkan efisiensi model dengan menghilangkan fitur yang kurang relevan. Model dapat dilatih lebih cepat dan mungkin menghasilkan prediksi yang lebih akurat.

Dengan mempertimbangkan saran-saran ini, peneliti berikutnya dapat memperluas pengetahuan tentang aplikasi Neural Network dan metode optimasi dengan ACO dan BP dalam penelitian medis, khususnya dalam prediksi diabetes tahap awal.

## DAFTAR PUSTAKA

- Alić, B., Gurbeta, L., & Badnjević, A. (2017). Machine learning techniques for classification of diabetes and cardiovascular diseases. *2017 6th Mediterranean Conference on Embedded Computing, MECO 2017 - Including ECYPS 2017, Proceedings, June*. <https://doi.org/10.1109/MECO.2017.7977152>
- Amato, F., López, A., Peña-Méndez, E. M., Vaňhara, P., Hampl, A., & Havel, J. (2013). Artificial neural networks in medical diagnosis. *Journal of Applied Biomedicine, 11*(2), 47–58. <https://doi.org/10.2478/v10136-012-0031-x>
- Apriliah, W., Kurniawan, I., Baydhowi, M., & Haryati, T. (2021). Prediksi Kemungkinan Diabetes pada Tahap Awal Menggunakan Algoritma Klasifikasi Random Forest. *Sistemasi, 10*(1), 163. <https://doi.org/10.32520/stmsi.v10i1.1129>
- Bianchi, L., Dorigo, M., Gambardella, L. M., & Gutjahr, W. J. (2009). A survey on metaheuristics for stochastic combinatorial optimization. *Natural Computing, 8*(2), 239–287. <https://doi.org/10.1007/s11047-008-9098-4>
- Blum, C. (2005). Ant colony optimization: Introduction and recent trends. *Physics of Life Reviews, 2*(4), 353–373. <https://doi.org/10.1016/j.plrev.2005.10.001>
- Blum, C., & López-Ibáñez, M. (2011). Ant Colony Optimization. In *The Industrial Electronics Handbook - Five Volume Set* (Issue January 2004). <https://doi.org/10.4249/scholarpedia.1461>
- Blum, C., & Socha, K. (2005). Training feed-forward neural networks with ant colony optimization: An application to pattern classification. *Proceedings - HIS 2005: Fifth International Conference on Hybrid Intelligent Systems, 2005*, 233–238. <https://doi.org/10.1109/ICHIS.2005.104>
- Dorigo, M., Maniezzo, V., & Colomi, A. (1996). Ant system: Optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 26*(1), 29–41. <https://doi.org/10.1109/3477.484436>
- Dorigo, M., & Stützle, T. (2004). *Ant Colony Optimization*. MIT Press <https://doi.org/10.7551/mitpress/1290.001.0001>
- ERGÜN, Ö. N., & O.İLHAN, H. (2021). Early Stage Diabetes Prediction Using Machine Learning Methods. *European Journal of Science and Technology, 29*, 52–57. <https://doi.org/10.31590/ejosat.1015816>
- Gao, W. (2011). Financial Data Forecasting by Evolutionary Neural. *Artificial Intelligence and Computational Intelligence, 262–269*.

- Gori, M., & Tesi, A. (1992). On the problem of local minima in backpropagation. In *IEEE Transactions on Pattern Analysis and Machine Intelligence* (Vol. 14, Issue 1, pp. 76–86). <https://doi.org/10.1109/34.107014>
- Gouri, A., & Dekaken, A. (2013). *Epigenetic pathways in type 2 diabetes and its complications*. 1(1), 1–5. <https://doi.org/10.3823/1400>
- Heaton, J. (2018). Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning. *Genetic Programming and Evolvable Machines*, 19(1–2), 305–307. <https://doi.org/10.1007/s10710-017-9314-z>
- Hecht-Nielsen, R. (1989). *Theory of the backpropagation neural network*. 593–605. <https://doi.org/10.1109/ijcnn.1989.118638>
- International Diabetes Federation. (2021). IDF Diabetes Atlas (10th ed.). International Diabetes Federation. <https://www.diabetesatlas.org>
- Karegowda, A. G., Manjunath, A. S., & Jayaram, M. A. (2011). Application of Genetic Algorithm Optimized Neural Network Connection Weights for Medical. *Soft Computing*, 2(2), 15–23.
- Maiyuriska, R. (2022). Penerapan Jaringan Syaraf Tiruan dengan Algoritma Backpropagation dalam Memprediksi Hasil Panen Gabah Padi. *Jurnal Informatika Ekonomi Bisnis*, 4, 28–33. <https://doi.org/10.37034/infec.v4i1.115>
- Marco, D. (2004). An Introduction to Ant Colony Optimization and Swarm Intelligence, . *Eight International Conference on Parallel Problem Solving From Nature*, (PPSN VIII)(May 2006).
- Mavrovouniotis, M., & Yang, S. (2013). Evolving neural networks using ant colony optimization with pheromone trail limits. *2013 13th UK Workshop on Computational Intelligence, UKCI 2013, Cci*, 16–23. <https://doi.org/10.1109/UKCI.2013.6651282>
- Nugroho, P. S., Tianingrum, N. A., Sunarti, S., Rachman, A., Fahrurrodzi, D. S., & Amiruddin, R. (2020). Predictor risk of diabetes mellitus in Indonesia, based on national health survey. *Malaysian Journal of Medicine and Health Sciences*, 16(1), 126–130.
- Of, D., & Mellitus, D. (2014). Diagnosis and classification of diabetes mellitus. *Diabetes Care*, 37(SUPPL.1), 81–90. <https://doi.org/10.2337/dc14-S081>
- Olabanjo, O., Wusu, A., & Mazzara, M. (2023). *Deep Unsupervised Machine Learning for Early Diabetes Risk Prediction using Ensemble Feature Selection and Deep Belief Neural Networks*. January, 1–10. <https://doi.org/10.20944/preprints202301.0208.v1>
- Permadi, J., Rhomadhona, H., & Aprianti, W. (2021). Perbandingan K-Nearest Neighbor Dan Backpropagation Neural Network Dalam Prediksi Resiko Diabetes Tahap Awal. *Kumpulan Jurnal Ilmu Komputer (KLIK)*, 08(3), 352–365.

- Qowiyyum, E. (2021). Pemberdayaan Masyarakat Melalui Program Indonesia Sehat dengan Pendekatan Keluarga ( PIS-PK ) untuk Meningkatkan Akses Pelayanan Kesehatan di Wilayah Puskesmas (Studi Kasus Puskesmas Pulo LOR, Kecamatan Jombang, Kabupaten Jombang). *Publika*, 9, 121–226.
- Ravi, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., & Yang, G. Z. (2017). Deep Learning for Health Informatics. *IEEE Journal of Biomedical and Health Informatics*, 21(1), 4–21. <https://doi.org/10.1109/JBHI.2016.2636665>
- Rumelhart, D. E., Hintont, G. E., & Williams, R. J. (1986). Learning Representations by Back-Propagating Errors. *Cognitive Modeling*, 2, 3–6. <https://doi.org/10.7551/mitpress/1888.003.0013>
- Schmidhuber, J. (2015). Deep Learning in neural networks: An overview. *Neural Networks*, 61, 85–117. <https://doi.org/10.1016/j.neunet.2014.09.003>
- Wang, H., Niu, D., & Cai, C. (2010). Evaluation of enterprise ERP system based on neural network optimized by ant colony. *Lecture Notes in Electrical Engineering*, 67 LNEE, 275–282. [https://doi.org/10.1007/978-3-642-12990-2\\_31](https://doi.org/10.1007/978-3-642-12990-2_31)
- Whitley, D., Starkweather, T., & Bogart, C. (1990). Genetic algorithms and neural networks: optimizing connections and connectivity. *Parallel Computing*, 14(3), 347–361. [https://doi.org/10.1016/0167-8191\(90\)90086-O](https://doi.org/10.1016/0167-8191(90)90086-O)
- WHO, Panos, & Loke, A. (n.d.). Diabetes. Retrieved May 23, 2023, from [https://www.who.int/health-topics/diabetes#tab=tab\\_1](https://www.who.int/health-topics/diabetes#tab=tab_1)
- Yang, W., Dall, T. M., Beronjia, K., Lin, J., Semilla, A. P., Chakrabarti, R., Hogan, P. F., & Petersen, M. P. (2018). Economic costs of diabetes in the U.S. in 2017. *Diabetes Care*, 41(5), 917–928. <https://doi.org/10.2337/dci18-0007>
- Zhang, G. P. (2000). Neural networks for classification: A survey. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, 30(4), 451–462. <https://doi.org/10.1109/5326.897072>