



Direktorat Riset dan Pengabdian Masyarakat Direktorat Jenderal Riset dan Pengembangan
Kementerian Riset, Teknologi, dan Pendidikan Tinggi
Gedung BPPT II Lantai 19, Jl. MH. Thamrin No. 8 Jakarta Pusat
<https://simlitabmas.ristekdikti.go.id/>

PROTEKSI ISI LAPORAN AKHIR PENELITIAN

Dilarang menyalin, menyimpan, memperbanyak sebagian atau seluruh isi laporan ini dalam bentuk apapun kecuali oleh peneliti dan pengelola administrasi penelitian

LAPORAN AKHIR PENELITIAN

ID Proposal: 633abdba-66f6-498c-a544-80251cca24c3
laporan akhir Penelitian: tahun ke-1 dari 1 tahun

1. IDENTITAS PENELITIAN

A. JUDUL PENELITIAN

Prediksi Prevalensi Diabetes Tipe 2 menggunakan Artificial Neural Network

B. BIDANG, TEMA, TOPIK, DAN RUMPUN BIDANG ILMU

Bidang Fokus RIRN / Bidang Unggulan Perguruan Tinggi	Tema	Topik (jika ada)	Rumpun Bidang Ilmu
Kesehatan	-		Bidang Kesehatan Umum Lain Yang Belum Tercantum

C. KATEGORI, SKEMA, SBK, TARGET TKT DAN LAMA PENELITIAN

Kategori (Kompetitif Nasional/ Desentralisasi/ Penugasan)	Skema Penelitian	Strata (Dasar/ Terapan/ Pengembangan)	SBK (Dasar, Terapan, Pengembangan)	Target Akhir TKT	Lama Penelitian (Tahun)
Penelitian Kompetitif Nasional			SBK Riset Pembinaan/ Kapasitas	2	1

2. IDENTITAS PENGUSUL

Nama (Peran)	Perguruan Tinggi/ Institusi	Program Studi/ Bagian	Bidang Tugas	ID Sinta	H-Index
EVA FIRDAYANTI BISONO - Ketua Pengusul	Institut Ilmu Kesehatan Bhakti Wiyata	Rekam Medis dan Informasi Kesehatan	Menyusun proposal penelitian, mengkoordinir pengumpulan data, menyusun luaran penelitian	6726467	0

	Kediri				
KRISNITA DWI JAYANTI - Anggota Pengusul	Institut Ilmu Kesehatan Bhakti Wiyata Kediri	Rekam Medis dan Informasi Kesehatan	Membantu ketua dalam penyusunan laporan penelitian, melakukan pengumpulan data, mengkoordinasikan analisis data penelitian	5987588	0
RATNA FRENTY NURKHALIM - Anggota Pengusul	Institut Ilmu Kesehatan Bhakti Wiyata Kediri	Rekam Medis dan Informasi Kesehatan	Menyusun luaran penelitian, merancang anggaran penelitian, menyusun laporan penelitian	5990773	0

3. MITRA KERJASAMA PENELITIAN (JIKA ADA)

Pelaksanaan penelitian dapat melibatkan mitra kerjasama, yaitu mitra kerjasama dalam melaksanakan penelitian, mitra sebagai calon pengguna hasil penelitian, atau mitra investor

Mitra	Nama Mitra
-------	------------

4. LUARAN DAN TARGET CAPAIAN

Luaran Wajib

Tahun Luaran	Jenis Luaran	Status target capaian (accepted, published, terdaftar atau granted, atau status lainnya)	Keterangan (url dan nama jurnal, penerbit, url paten, keterangan sejenis lainnya)
1	Artikel di Jurnal Nasional Terakreditasi Peringkat 1-6	Accepted	https://jmiki.apfirmik.or.id/jmiki
1	Artikel di Jurnal Nasional Terakreditasi Peringkat 1-6	Submitted	https://jmiki.apfirmik.or.id/jmiki
1	Feasibility Study		

Luaran Tambahan

Tahun Luaran	Jenis Luaran	Status target capaian (accepted, published, terdaftar atau granted, atau status lainnya)	Keterangan (url dan nama jurnal, penerbit, url paten, keterangan sejenis lainnya)
--------------	--------------	--	---

5. ANGGARAN

Rencana anggaran biaya penelitian mengacu pada PMK yang berlaku dengan besaran minimum dan

maksimum sebagaimana diatur pada buku Panduan Penelitian dan Pengabdian kepada Masyarakat

Total RAB 1 Tahun Rp. 0

Tahun 1 Total Rp. 0

Jenis Pembelanjaan	Komponen	Item	Satuan	Vol.	Biaya Satuan	Total
--------------------	----------	------	--------	------	--------------	-------

Tahun 2 Total Rp. 0

Jenis Pembelanjaan	Komponen	Item	Satuan	Vol.	Biaya Satuan	Total
--------------------	----------	------	--------	------	--------------	-------

Tahun 3 Total Rp. 0

Jenis Pembelanjaan	Komponen	Item	Satuan	Vol.	Biaya Satuan	Total
--------------------	----------	------	--------	------	--------------	-------

6. KEMAJUAN PENELITIAN

A. RINGKASAN

Perkembangan kejadian jumlah DM tipe 2 di kabupaten Kediri tahun 2021 meningkat sebesar 2% dibandingkan tahun 2020. Kondisi ini mengharuskan dilakukannya forecasting untuk memprediksi prevalensi DM di masa yang akan datang. Forecasting digunakan sebagai langkah mitigasi untuk dalam pengendalian kejadian DM tipe 2 dengan cara memprediksi kejadian DM tipe 2 di masa depan. Metode forecasting yang digunakan dalam penelitian ini adalah Artificial Neural Network (ANN). Metode ANN merupakan metode berbasis kecerdasan buatan (Artificial Intelligence) dengan menggunakan konsep yang terdiri dari sejumlah besar unit pemrosesan sederhana yang disebut neuron atau node, yang terhubung satu sama lain melalui koneksi bobot. Setiap neuron menerima masukan dari neuron lainnya, menghitung hasilnya, dan kemudian mengirimkan hasilnya ke neuron

lainnya. Koneksi antara neuron diberi bobot yang menentukan kekuatan hubungan antara mereka. ANN mempelajari pola dari data yang disajikan padanya melalui algoritma pembelajaran, dan kemudian dapat digunakan untuk mengklasifikasikan atau memprediksi data baru. Metode ANN memiliki beberapa keunggulan dibandingkan metode klasik seperti Autoregressive Integrated Moving Average (ARIMA). Keunggulan dari metode ini adalah mampu menangkap informasi yang bersifat nonlinier pada data time series dan tidak memerlukan asumsi, dan memiliki akurasi yang lebih baik dibandingkan metode ARIMA. Berdasarkan pemaparan sebelumnya, penelitian ini bertujuan untuk melakukan prediksi prevalensi penyakit DM tipe 2 menggunakan Artificial Neural Network (ANN). Penelitian ini menggunakan data rekam medis elektronik Rumah Sakit HVA Toeloengredjo Pare

Kabupaten Kediri. Data merupakan data sekunder yang didapatkan dari database rumah sakit mulai periode Januari 2019 sampai 2022. Hasil analisis menggunakan metode CART memberikan kesimpulan bahwa variabel morbiditas, urbanitas, komplikasi, dan gender merupakan variabel yang dapat digunakan untuk memprediksi LOS pada pasien DM tipe 2. Variabel kesakitan merupakan variabel yang mempunyai tingkat kepentingan paling tinggi dibandingkan variabel lainnya. Model CART memiliki akurasi sebesar 0,6574 (65,74%), dan Skor F1 sebesar 0,76. Luaran penelitian ini adalah dapat publish pada jurnal nasional

yang terakreditasi

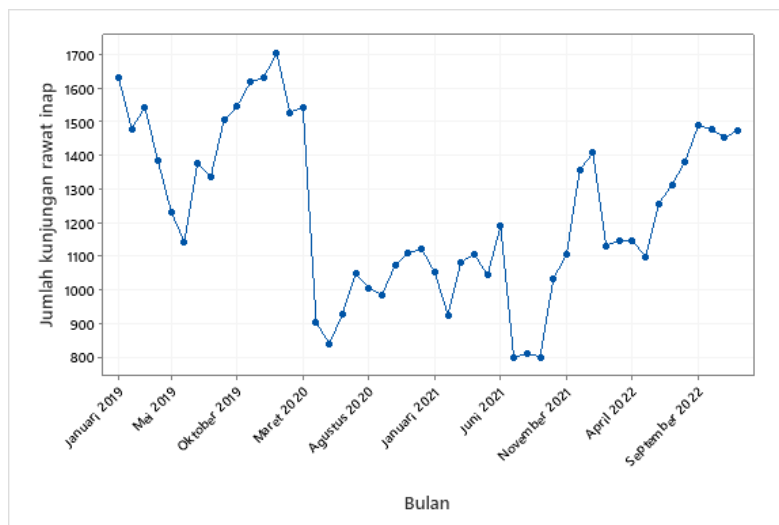
B. KATA KUNCI

Artificial Neural Network; Diabetes Mellitus tipe 2; Prevalensi

Pengisian poin C sampai dengan poin H mengikuti template berikut dan tidak dibatasi jumlah kata atau halaman namun disarankan ringkas mungkin. Dilarang menghapus/memodifikasi template ataupun menghapus penjelasan di setiap poin.

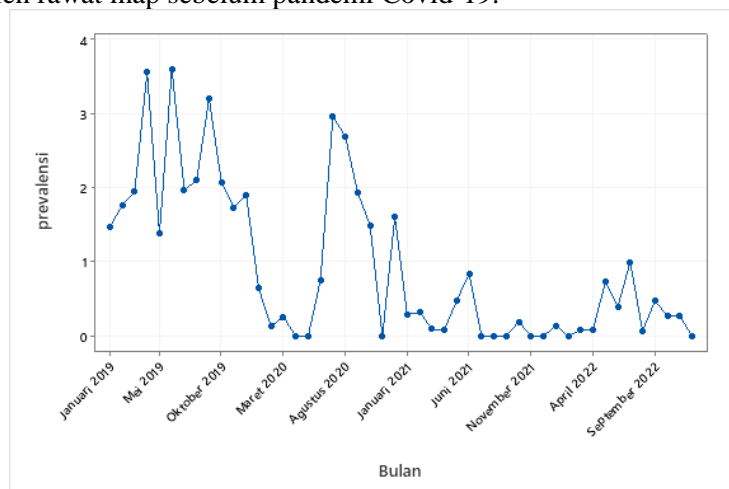
C. HASIL PELAKSANAAN PENELITIAN: Tuliskan secara ringkas hasil pelaksanaan penelitian yang telah dicapai sesuai tahun pelaksanaan penelitian. Penyajian meliputi data, hasil analisis, dan capaian luaran (wajib dan atau tambahan). Seluruh hasil atau capaian yang dilaporkan harus berkaitan dengan tahapan pelaksanaan penelitian sebagaimana direncanakan pada proposal. Penyajian data dapat berupa gambar, tabel, grafik, dan sejenisnya, serta analisis didukung dengan sumber pustaka primer yang relevan dan terkini.

Data kunjungan pasien rawat inap selama periode tahun 2019 sampai 2022 mengalami fluktuasi. Time series plot kunjungan pasien di rumah sakit selama periode tersebut disajikan pada gambar 1 di bawah ini.



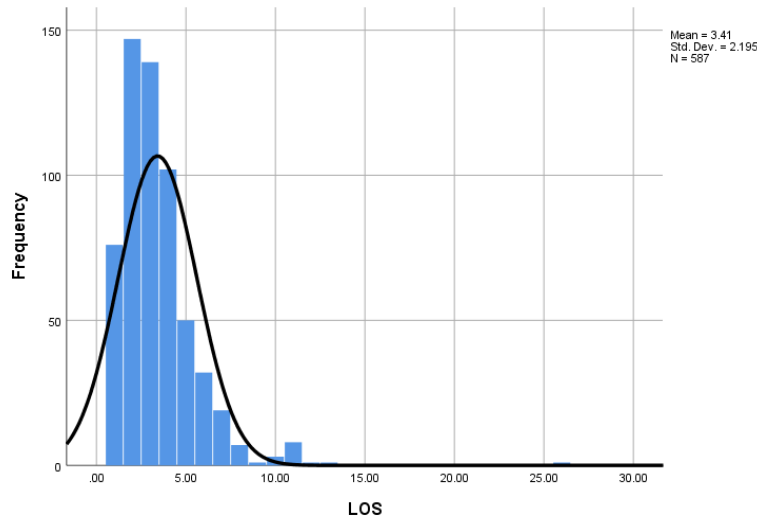
Gambar 1. Plot Time series kunjungan pasien rawat inap 2019 - 2022

Gambar 1 memberikan informasi bahwa kunjungan pasien rawat inap mengalami penurunan signifikan pada bulan maret tahun 2020. Kondisi ini dikarenakan pada saat tersebut terjadi awal mula pandemi Covid-19 di Indonesia. Pada bulan juni sampai Juli sampai September 2021, jumlah pasien rawat inap juga mengalami penurunan dengan rata-rata sekitar 800 pasien setiap bulan. Pada tahun 2021, jumlah pasien rawat inap sudah mengalami kenaikan dan ada kecenderungan jumlahnya sama dengan jumlah pasien rawat inap sebelum pandemi Covid-19.



Gambar 2. Plot Time series prevalensi DM tipe 2 periode 2019 - 2022

Prevalensi pasien DM tipe 2 dihitung berdasarkan formula seperti yang dijelaskan pada desain penelitian yaitu $\text{Prevalensi} = \frac{\text{kasus lama} + \text{kasus baru}}{\text{jumlah pasien dalam periode tersebut}} \times 100\%$. Hasil perhitungan prevalensi pasien DM tipe 2 disajikan pada tabel 1. Gambar 2 memberikan visualisasi trend prevalensi pada periode 2019-2022. Berdasarkan gambar 2, dapat dilihat bahwa sebelum tahun 2020, prevalensi DM tipe 2 di rumah sakit cenderung lebih tinggi dibandingkan setelah tahun 2020. Pada tahun 2021 bahkan ditemukan pada bulan Agustus, November, dan Desember prevalensi DM tipe 2 sebesar 0.



Gambar 3. Histogram Length of stay (LoS) pasien DM tipe 2 periode 2019 - 2022

Hasil preprocessing data, memberikan informasi jumlah kunjungan rawat inap pasien DM tipe 2 selama periode tahun 2019 – 2022 sebanyak 578. Histogram pada gambar 1 memberikan informasi tentang sebaran lama dirawat (Length of Stay) pasien DM tipe 2 di rumah sakit. Berdasarkan gambar tersebut dapat dilihat bahwa LoS pasien DM tipe 2 paling banyak disekitar 2-3 hari. Histogram tersebut juga memberikan informasi bahwa juga terdapat pasien DM tipe 2 dengan LoS lebih dari 10 hari.

Tabel 1. Statistik Deskriptif LoS pasien DM tipe 2.

Parameter	Nilai
Mean	3,14
Median	3
Minimum	1
Maksimum	26
Standard Deviasi	2,19

Berdasarkan informasi pada tabel 1 di atas, dapat diketahui bahwa rata-rata lama dirawat (Length of stay) pasien DM tipe 2 sebesar 3,14 hari. Lama dirawat paling cepat adalah 1 hari, dan berdasarkan informasi pada tabel tersebut, juga terdapat pasien dengan lama dirawat selama 26 hari.

Tabel 2. Distribusi frekuensi pasien DM tipe 2 berdasarkan jenis kelamin

Jenis Kelamin	Frekuensi	Persentase
Perempuan	366	62.4
Laki-laki	221	37.6

Total	587	100
-------	-----	-----

Karakteristik pasien DM tipe 2 berdasarkan jenis kelamin disajikan pada tabel 2. Pasien DM tipe 2 yang berjenis kelamin perempuan sebanyak 366 pasien atau sekitar 62,4% dari total pasien DM tipe 2. Pasien laki-laki sebanyak 221 pasien (37,6%). Informasi ini memberikan gambaran bahwa perempuan cenderung lebih beresiko terkena DM tipe 2. Hasil ini selaras dengan penelitian yang dilakukan oleh Wulandari pada tahun 2020 yang menghasilkan kesimpulan bahwa wanita memiliki risiko 3-7 kali lebih banyak terkena Diabetes Melitus tipe 2 (15). Penelitian lain juga memberikan hasil yang konsisten. Pada tahun 2021, Gunawan dan Rahmawati melakukan penelitian dan menghasilkan kesimpulan yaitu bahwa frekuensi pasien diabetes tipe 2 perempuan lebih banyak dibandingkan laki-laki (16).

Tabel 3. Distribusi frekuensi pasien DM tipe 2 berdasarkan usia

Usia	Frekuensi	Persentase
5-14 TH	2	0.3
15-24 TH	4	0.7
25-44 TH	45	7.7
45-64 TH	377	64.2
>= 65 TH	159	27.1
Total	587	100

Pada penelitian ini, diketahui bahwa pasien DM tipe paling banyak memiliki rentang usia 45-64 tahun. Pada rentang usia ini terdapat 377 pasien DM tipe 2 atau sekitar 64,2% dari total seluruh data penelitian. Informasi ini sesuai dengan penelitian yang lain yang memberikan kesimpulan bahwa memiliki risiko 11,183 kali lebih besar untuk terkena kasus Diabetes Melitus Tipe 2 (17).

Tabel 4. Distribusi frekuensi pasien DM tipe 2 berdasarkan adanya penyakit penyerta

Status morbiditas	Frekuensi	Persentase
Tidak memiliki penyerta	224	38.2
Memiliki penyerta	363	61.8
Total	587	100

Tabel 4 menjelaskan distribusi frekuensi adanya penyakit penyerta pada pasien DM tipe 2. Berdasarkan informasi pada tabel tersebut, diketahui bahwa 363 pasien DM tipe 2 memiliki penyakit penyerta. Sedangkan pasien DM tipe 2 yang tidak memiliki penyakit penyerta sebanyak 224 pasien atau sekitar 38,2%.

Tabel 5. Distribusi frekuensi pasien DM tipe 2 berdasarkan adanya komplikasi

Status Komplikasi	Frekuensi	Persentase
Komplikasi	229	39
Tidak	358	61
Total	587	100

Penyakit DM tipe 2 dapat menyebabkan komplikasi terhadap penyakit lain. Tabel 5 memberikan informasi jumlah pasien yang mengalami komplikasi penyakit lain. Dari total 578 pasien DM tipe 2, 229 pasien (39%) mengalami komplikasi penyakit lain. Jenis komplikasi yang diderita pasien DM tipe 2 disajikan pada tabel 6 di bawah ini.

Tabel 6. Distribusi frekuensi komplikasi yang diderita pasien DM tipe 2

Komplikasi yang diderita	Frekuensi	Persentase
Type 2 diabetes mellitus with hyperosmolarity	16	2.7
Type 2 diabetes mellitus with ketoacidosis	14	2.4
Type 2 diabetes mellitus with kidney complications	19	3.2
Type 2 diabetes mellitus with ophthalmic complications	1	0.2
Type 2 diabetes mellitus with neurological complications	11	1.9
Type 2 diabetes mellitus with circulatory complications	115	19.6
Type 2 diabetes mellitus with other specified complications	3	0.5
Type 2 diabetes mellitus with multiple complications	21	3.6
Type 2 diabetes mellitus with with unspecified complications	29	4.9
Type 2 diabetes mellitus without complication	358	61
Total	587	100

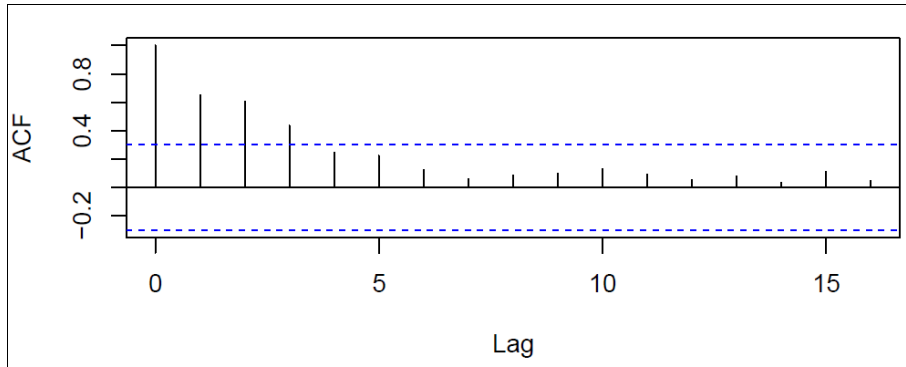
Type 2 diabetes mellitus with circulatory complications merupakan komplikasi yang paling banya diderita pada pasien penelitian ini. Pasien yang mengalami komplikasi ini sebanyak 115 pasien. Penyakit DM tipe 2 juga berdampak kepada komplikasi terjadinya penyakit ginjal. Berdasarkan tabel 6 diketahui terdapat 19 pasien DM tipe 2 yang mengalami komplikasi dengan penyakit ginjal.

Tabel 7. Distribusi frekuensi pasien DM tipe 2 berdasarkan asal pasien

Asal Pasien	Frekuensi	Persentase
IGD	539	91.8
Poli	48	8.2
Total	587	100

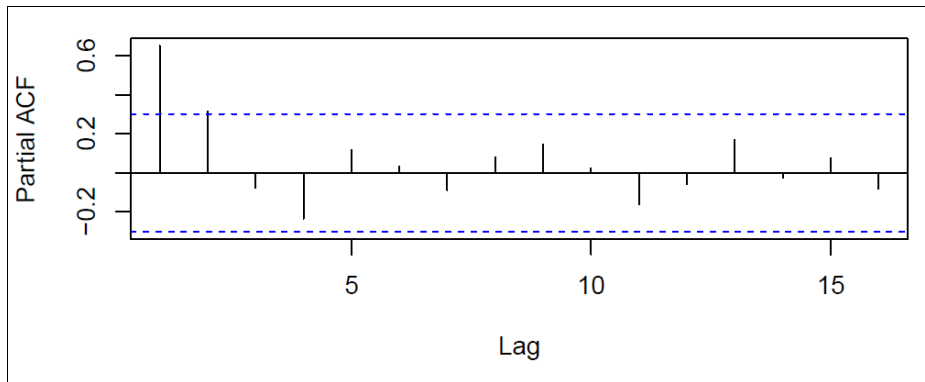
Pasien rawat inap DM tipe 2 pada penelitian ini berasal dari IGD dan Poli. Pasien rawat inap DM tipe 2 yang berasal dari poli sebanyak 48 pasien (8,2%). Sedangkan pasien rawat inap yang berasal dari IGD, ada sebanyak 539 (91,8%) dari total pasien DM tipe 2.

Analisis statistika tahap selanjutnya adalah melakukan analisis prediksi prevalensi pasien DM tipe 2 selama periode 2019 sampai 2022. Metode yang digunakan adalah Neural network dengan algoritma feedforward. Langkah awal dalam pemodelan ANN adalah identifikasi struktur data deret waktu yang akan digunakan sebagai input layer pada ANN. Identifikasi dilakukan dengan melihat plot ACF dan plot PACF dari data prevalensi DM tipe 2 tahun 2019 – 2022.



Gambar 4. Plot ACF prevalensi DM tipe 2 tahun 2019-2022

Gambar 4 memberikan informasi tentang plot ACF prevalensi DM tipe 2 selama periode 2019-2022. Berdasarkan plot ACF tersebut, diketahui pola ACF turun secara lambat (eksponensial). Identifikasi model ARIMA juga mempertimbangkan plot PACF dari prevalensi pasien DM tipe 2.



Gambar 5. Plot PACF prevalensi DM tipe 2 tahun 2019-2022

Gambar 5 memberikan informasi tentang plot PACF prevalensi DM tipe 2 selama periode 2019-2022. Identifikasi model ARIMA berdasarkan Montgomery tahun 2015, apabila ACF turun secara lambat dan plot PACF *cut off* pada lag ke -p, maka model ARIMA yang dapat dibentuk adalah AR(p). Berdasarkan plot PACF tersebut, diketahui pola PACF *cut off* setelah lag 2, maka model ARIMA yang dapat dibangun adalah model AR(2) atau ARIMA (2,0,0).

Model AR(2) memiliki makna bahwa data pada pengamatan ke n, dipengaruhi oleh data ke n-1 dan n-2. Model AR(2) ini yang mendasari struktur input data pada model ANN. Struktur ANN dalam penelitian ini dijelaskan pada gambar 6.

Gambar 6. Struktur data time series untuk input model ANN.

Tabel 8. Prediksi prevalensi pada data testing dengan variasi jumlah neuron

Bulan	Aktual Prevalensi	Jumlah Neuron pada hidden layer								
		2	3	4	5	6	7	8	9	10
Juli	0.9901	0.7856	0.7888	0.8054	0.7807	0.7815	0.7887	0.7855	0.7948	0.7959
Agustus	0.0724	0.5693	0.5791	0.5890	0.5821	0.5810	0.5801	0.5820	0.5842	0.5873

September	0.4698	0.8223	0.8270	0.8549	0.8136	0.8146	0.8266	0.8225	0.8383	0.8392
Oktober	0.2708	0.6716	0.6864	0.7103	0.6924	0.6901	0.6928	0.6912	0.6973	0.7020
November	0.2751	0.8455	0.8508	0.8863	0.8340	0.8350	0.8500	0.8453	0.8665	0.8666
Desember	0.0000	0.7346	0.7490	0.7899	0.7594	0.7571	0.7626	0.7585	0.7697	0.7732

Pemilihan jumlah neuron pada hidden layer didasarkan pada hasil prediksi model neural network menggunakan data testing. Pada penelitian ini data testing adalah angka prevalensi DM tipe 2 bulan juli 2022 – desember 2022. Tabel 8 di atas memberikan informasi prediksi angka prevalensi bulan juli 2022 sampai desember 2022. Hasil prediksi tersebut akan digunakan sebagai dasar perhitungan RMSE untuk menentukan jumlah neuroan yang paling optimal.

Tabel 9. Nilai RMSE model ANN dengan neuron yang berbeda pada hidden layer

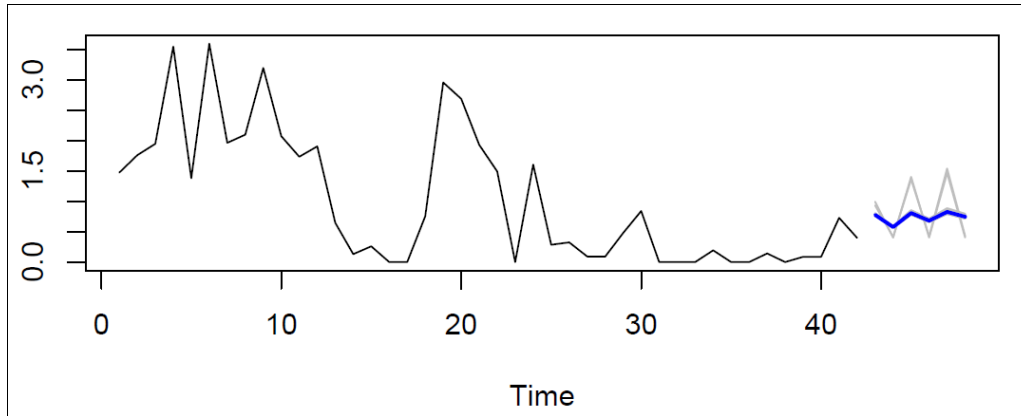
RMSE	Jumlah Neuron pada hidden layer								
	2	3	4	5	6	7	8	9	10
RMSE	0.4912	0.6414	0.8057	0.9270	1.0691	1.2165	1.3576	1.5086	1.6533

Pada tabel 9 dapat lihat nilai RMSE hasil prediksi prevalensi dari model neural network dimulai dari jumlah neuron sebanyak 2 sampai 10. Berdasarkan tabel tersebut dapat diketahui bahwa nilai RMSE yang paling kecil didapatkan ketika menggunakan 2 neuroan pada hidden layer. Prediksi prevalensi bulan januari 2023 sampai juni 2023 dilakukan menggunakan model Neural network dengan 2 neuron pada hidden layer.

Tabel 9. Prediksi prevalensi menggunakan 2 neuron pada hidden layer.

Bulan	Prediksi
Januari 2023	0.776
Februari 2023	0.576
Maret 2023	0.806
April 2023	0.681
Mei 2023	0.825
Juni 2023	0.745

Prediksi prevalensi DM tipe 2 dengan model ANN yang terdiri dari dua neuron pada hidden layer disajikan pada tabel 9. Berdasarkan tabel 9, dapat diketahui hasil prediksi berfluktuasi. Pada bula januari prediksi angka prevalensi sebesar 0,776. Sedangkan pada bulan juni 2023, prediksi prevalensi sebesar 0,745. Time series plot prediksi prevalensi januari sampai juni 2023 disajikan pada gambar 6 di bawah ini.






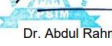
Gambar 6. Time series plot prediksi prevalensi Januari – Juni 2023.

Berdasarkan gambar 6 di atas, dapat dilihat bahwa hasil prediksi cenderung stagnan., dan belum terlihat pola trend menurun. Dalam data tersebut juga perlu diperhatikan bahwa pada bulan tertentu, terdapat angka prevalensinya sebesar 0. Tentunya kondisi ini dapat berpengaruh terhadap hasil prediksi model ANN.

D. STATUS LUARAN: Tuliskan jenis, identitas dan status ketercapaian setiap luaran wajib dan luaran tambahan (jika ada) yang dijanjikan. Jenis luaran dapat berupa publikasi, perolehan kekayaan intelektual, hasil pengujian atau luaran lainnya yang telah dijanjikan pada proposal. Uraian status luaran harus didukung dengan bukti kemajuan ketercapaian luaran sesuai dengan luaran yang dijanjikan. Lengkapi isian jenis luaran yang dijanjikan serta mengunggah bukti dokumen ketercapaian luaran wajib dan luaran tambahan melalui BIMA.

Luaran dan target capaian yang dijanjikan

a. Luaran Wajib			
No	Jenis Luaran	Status target capaian	Keterangan
1	Artikel di Jurnal Nasional terakreditasi peringkat 1-6. JUITA: Jurnal Informatika ISSN : 2579-8901 (online) 2086-9398 (print)	Accepted	Accepted (LOA Telah Terbit) <div style="text-align: center;">  JUITA JURNAL INFORMATIKA – 2086-9398 (p-ISSN); 2579-8901(e-ISSN) <small>Jl. Raya Dukuwaluh P.O. Box 202 Purwokerto 53182 Telp. (0281) 636751, 630463 Fax. (0281) 637239</small> </div> <hr/> <p style="text-align: center;">ACCEPTANCE LETTER</p> <p>Dear Mr./Mrs./Ms. Eva Firdayanti Bisono</p> <p>It is my great pleasure to inform you that your contribution to JUITA: Jurnal Informatika,</p> <p>Reference No. : 19564</p> <p>Title : LENGTH OF STAY (LOS) PREDICTION OF TYPE 2 DIABETES MELLITUS USING CLASSIFICATION AND REGRESSION TREE (CART)</p> <p>Author(s) : Eva Firdayanti Bisono, Jerhi Wahyu Fernanda, Ratna Frenty Nurkhalim, Krisnita Dwi Jayanti</p> <p>has been accepted for the publication on Vol. 12 No. 1, Mei 2024.</p> <p>We would like you to refer to JUITA publications especially your own works in your future research. We believe it will help you promote your article and enhance its visibility.</p>
b. Luaran tambahan			

No	Jenis Luaran	Status target capaian	Keterangan
1	Buku Ajar	Publish	<p>Penerbit</p>  <p>PENERBIT BUKU YAYASAN PENDIDIKAN DAN SOSIAL INDONESIA MAJU BCP 2 Blok E. 18 No.14 Desa Ranjeng, Kec. Ciruas, Kab. Serang Wa. 08159516818. Email: ypsimbanten@gmail.com, http://ypsimbanten.com</p> <hr/> <p>KWITANSI</p> <p>Nomor : 052/PEN/XII/YPSIM/2023 Telah diterima dari : Krisnita Dwi Jayanti.,S.KM.M.Epid Uang Sejumlah : <i>Delapan Ratus Lima Puluh Ribu Rupiah</i> Untuk pembayaran : Penerbitan Buku</p> <p>Serang, 10 Desember 2023 Pimpinan</p> <div style="border: 1px solid orange; padding: 2px; display: inline-block;">Rp. 850.000,-</div>   Dr. Abdul Rahman H., M.T, CT, CHCP, C.PI

E. PERAN MITRA: Tuliskan realisasi kerjasama dan kontribusi Mitra baik *in-kind* maupun *in-cash* (untuk Penelitian Terapan, Penelitian Pengembangan, PTUPT, PPUPT serta KRUP). Bukti pendukung realisasi kerjasama dan realisasi kontribusi mitra dilaporkan sesuai dengan kondisi yang sebenarnya. Bukti dokumen realisasi kerjasama dengan Mitra diunggah melalui BIMA.

.....

.....

.....

.....

.....

F. KENDALA PELAKSANAAN PENELITIAN: Tuliskan kesulitan atau hambatan yang dihadapi selama melakukan penelitian dan mencapai luaran yang dijanjikan, termasuk penjelasan jika pelaksanaan penelitian dan luaran penelitian tidak sesuai dengan yang direncanakan atau dijanjikan.

Kendala yang dihadapi ketika penelitian adalah pada saat pengumpulan data. Diperlukan beberapa kali kunjungan ke rumah sakit untuk memaparkan kepada jajaran direktur mengenai penggunaan data, aspek kerahasiaan data dan manfaat penelitian

G. RENCANA TAHAPAN SELANJUTNYA: Tuliskan dan uraikan rencana penelitian di tahun berikutnya berdasarkan indikator luaran yang telah dicapai, rencana realisasi luaran wajib yang dijanjikan dan tambahan (jika ada) di tahun berikutnya serta *roadmap* penelitian keseluruhan. Pada bagian ini diperbolehkan untuk melengkapi penjelasan dari setiap tahapan dalam metoda yang akan direncanakan termasuk jadwal berkaitan dengan strategi untuk mencapai luaran seperti yang telah dijanjikan dalam proposal. Jika diperlukan, penjelasan dapat juga dilengkapi dengan gambar, tabel, diagram, serta pustaka yang relevan. Pada bagian ini dapat dituliskan rencana penyelesaian target yang belum tercapai.

Rencana tahapan selanjutnya untuk penelitian pada tahun selanjutnya adalah mengembangkan model statistik spasial dan machine learning untuk mengetahui pola penyebaran Diabetes Mellitus Tipe 2 di kota Kediri menggunakan sumber data survey maupun data rekam medis

H. DAFTAR PUSTAKA: Penyusunan Daftar Pustaka berdasarkan sistem nomor sesuai dengan urutan pengutipan. Hanya pustaka yang disitasi pada laporan akhir yang dicantumkan dalam Daftar Pustaka.

1. Dinkes Jatim. Profil Kesehatan Dinas Kesehatan Provinsi Jawa Timur 2021. Dinas Kesehat Provinsi Jawa Timur. 2021;1–149.
2. Yang E, Park HW, Choi YH, Kim J, Munkhdalai L, Musa I, et al. A simulation-based study on the comparison of statistical and time series forecasting methods for early detection of infectious disease outbreaks. *Int J Environ Res Public Health*. 2018;15(5).
3. Anwar MY, Lewnard JA, Parikh S, Pitzer VE. Time series analysis of malaria in Afghanistan: using ARIMA models to predict future trends in incidence. *Malar J*. 2016;15(1):1–10.
4. Xu B, Li J, Wang M. Epidemiological and time series analysis on the incidence and death of AIDS and HIV in China. *BMC Public Health*. 2020;20(1):1–10.
5. Shahid N, Rappon T, Berta W. Applications of artificial neural networks in health care organizational decision-making: A scoping review. *PLoS One*. 2019;14(2):1–22.
6. Wang KW, Deng C, Li JP, Zhang YY, Li XY, Wu MC. Hybrid methodology for tuberculosis incidence time-series forecasting based on ARIMA and a NAR neural network. *Epidemiol Infect*. 2017;145(6):1118–29.
7. Lou HR, Wang X, Gao Y, Zeng Q. Comparison of ARIMA model, DNN model and LSTM model in predicting disease burden of occupational pneumoconiosis in Tianjin, China. *BMC Public Health* [Internet]. 2022;22(1):1–15. Available from: <https://doi.org/10.1186/s12889-022-14642-3>
8. Zheng Y, Zhang X, Wang X, Wang K, Cui Y. Predictive study of tuberculosis incidence by time series method and Elman neural network in Kashgar, China. *BMJ Open*. 2021;11(1):1–8.
9. Tang K, Luo R, Zhang S. An Artificial Neural Network Algorithm for the Evaluation of Postoperative Rehabilitation of Patients. *J Healthc Eng*. 2021;2021.
10. Jerhi Wahyu Fernanda, Sidjabat FN. Prediksi Incidence Dengue Hemorrhagic Fever (Dhf) Menggunakan Jaringan Saraf Tiruan (Artifial Neural Network). *J Manaj Inf Kesehat Indones*. 2018;6(2):142.
11. Lutz CS, Huynh MP, Schroeder M, Anyatonwu S, Dahlgren FS, Danyluk G, et al. Applying infectious disease forecasting to public health: A path forward using influenza forecasting examples. *BMC Public Health*. 2019;19(1):1–12.
12. Rabelo L, Morais DA, Soares G. Forecasting daily Covid-19 cases in the world with a hybrid ARIMA and neural network model. 2020;126(January).
13. James A, Tripathi V. Time series data analysis and ARIMA modeling to forecast the short-term trajectory of the acceleration of fatalities in Brazil caused by the corona virus (COVID-19). *PeerJ*. 2021;9.
14. Fernanda JW, . P, Wisnaningsih ER. Prediksi Ketahanan Hidup Pasien Kanker Kolorektal Menggunakan Jaringan Saraf Tiruan (Artificial Neural Network). *J Manaj Inf Kesehat Indones*. 2018;6(1):46.
15. Wulandari IAT, Herawati S, Wande IN. Program Studi Sarjana Kedokteran dan Profesi Dokter, Fakultas Kedokteran Universitas Udayana 2 Departemen Patologi Klinik Fakultas Kedokteran Universitas Udayana Koresponding author: Ida Ayu Trisna Wulandari. *J Med Udayana*. 2020;9(1):71–5.

16. Gunawan S, Rahmawati R. Hubungan Usia, Jenis Kelamin dan Hipertensi dengan Kejadian Diabetes Mellitus Tipe 2 di Puskesmas Tugu Kecamatan Cimanggis Kota Depok Tahun 2019. *ARKESMAS (Arsip Kesehat Masyarakat)*. 2021;6(1):15–22.
17. Adri K, Arsin A, Thaha RM. Faktor Risiko Kasus Diabetes Mellitus Tipe 2 Dengan Ulkus Diabetik Di Rsud Kabupaten Sidrap. *J Kesehat Masy Marit*. 2020;3(1):101–8.

Ringkasan eksekutif maksimum 500 kata: memberikan gambaran umum tentang isi yang terkandung dalam dokumen studi kelayakan. Bagian ini merupakan ringkasan poin penting dari detail yang terkandung dalam keseluruhan dokumen studi kelayakan dan deskripsi singkat tentang produk dan/atau jasa yang dianggap sudah melalui tahapan kajian sebelumnya.

Ringkasan: Buku yang diusulkan bertujuan untuk memberikan wawasan komprehensif tentang pemanfaatan teknik data mining untuk menggali informasi berharga dari catatan medis pasien diabetes. Tujuannya adalah untuk meningkatkan hasil perawatan kesehatan, mengoptimalkan rencana pengobatan, dan berkontribusi pada penelitian diabetes melalui pendekatan berbasis data. Pasar untuk literatur terkait diabetes dan analitika data dalam bidang kesehatan berkembang pesat karena prevalensi diabetes yang semakin tinggi secara global. Terdapat permintaan yang signifikan untuk wawasan yang dapat diambil dari data medis untuk meningkatkan perawatan pasien, sehingga buku ini sangat relevan dan berpotensi menguntungkan. Pembaca utama buku ini ialah mahasiswa rekam medis dan informasi kesehatan untuk meningkatkan pemahaman mereka tentang manajemen diabetes melalui data mining. Buku ini akan mencakup berbagai topik, termasuk pengenalan tentang diabetes dan komplikasinya, dasar-dasar data mining, prapemrosesan data medis, teknik data mining yang relevan (mis. clustering, klasifikasi, asosiasi), studi kasus, dan aplikasi praktis. Isi akan disusun untuk memastikan pembelajaran yang progresif, memungkinkan pembaca untuk menerapkan konsep data mining dengan efektif. Studi kelayakan menunjukkan pandangan positif untuk pengembangan buku yang difokuskan pada data mining pada rekam medis pasien diabetes. Dengan pasar yang terus berkembang, audiens target yang terdefinisi dengan baik, konten yang tepat, dan sarana teknis dan keuangan untuk memproduksi buku, disarankan untuk melanjutkan proyek ini.

Pasar Produk/Layanan maksimum 500 kata: menjelaskan pasar yang ada untuk produk dan/atau jasa yang sedang dikembangkan. Peneliti sebaiknya memaparkan keunggulan-keunggulan kompetitif dan komparatif produk/jasa yang sedang dikembangkan, dengan membandingkannya dengan calon pesaing, mampu menawarkan nilai yang lebih besar kepada calon konsumen dari pada yang ditawarkan pesaing, serta keunikan-keunikan tertentu dari produk/hasil penelitiannya yang sulit ditiru produsen lain, perkiraan pasar yang bisa direbut.

Pasar Produk/Layanan: Produk dari penelitian ini adalah buku ajar metode data mining dalam bidang rekam medis dan informasi kesehatan. Perkembangan teknologi informasi yang menuntut penyimpanan data data pasien dalam bentuk elektronik, akan menghasilkan suatu data yang tersimpan dalam database SIMRS rumah sakit. Data ini semakin lama akan berkembang menjadi big data, sehingga sangat diperlukan proses pengolahan dan analisis untuk mendapatkan informasi yang akurat. Proses pengolahan dan analisis data dalam database akan memerlukan suatu metode yang tepat salah satunya metode data mining. Dalam buku ini akan dijelaskan proses pengolahan data yang dimulai import data, penggabungan data, dan filter data dari suatu dataset yang didapatkan dari database. Buku ini juga akan menjelaskan Teknik-teknik

visualisasi data yang bermanfaat dalam penyampaian informasi dari data seperti time series plot, diagram lingkaran, diagram batang, histogram, boxplot, dan Teknik visualisasi data yang lain. Buku ini juga akan menjelaskan metode klasifikasi khususnya decision tree dalam pemodelan data kesehatan. Pada kasus pengenalan trend penyakit, metode time series untuk pengenalan pola trend suatu data juga dijelaskan dalam buku ini.

Pertimbangan Teknologi/Sosial maks 500 kata: menjelaskan pertimbangan apa saja yang dibuat oleh peneliti terkait dengan aspek teknologi, lingkungan, sosial, dan hukum. Peneliti perlu menjelaskan bahwa teknologi atau solusi teknis yang diusulkan implementatif dan kompetitif, serta apakah saat ini mereka menguasai teknologi dan keahlian teknis yang diperlukan tersebut. Peneliti perlu memaparkan sumber dari teknologi yang dipakai, apakah dari internal atau eksternal, serta HKI dari teknologi-teknologi tersebut. Perlu dijelaskan apakah perlu mengembangkan teknologi baru, atau cukup menggunakan teknologi yang ada, serta kemungkinan untuk membeli teknologi yang sudah ada.

Pertimbangan Teknologi/Sosial : Data yang didapatkan dari rumah sakit terdiri dari data rawat inap pasien setiap tahun dalam periode 2019 sampai 2022. Untuk proses analisis statistik, diperlukan proses preprocessing data untuk menggabungkan data-data tersebut. Proses penggabungan dan pengolahan data awal menggunakan software R. Software R merupakan software yang berfungsi sebagai tools untuk analisis statistika yang bersifat opensource. Software ini terdiri dari banyak library yang dapat digunakan untuk proses pengolahan dan analisis data statistika di bidang kesehatan, social, dan bidang lainnya. Software ini merupakan solusi alternative selain software SPSS, minitab, Stata dalam proses pengolahan dan analisis data. Tentunya dengan perkembangan teknologi dan data ke arah big data, software ini akan sangat dibutuhkan dalam proses pengolahan dan analisis data karena fleksibilitas yang dimiliki. Penelitian ini menggunakan metode Artificial Neural Network (ANN) untuk pemodelan dan prediksi angka prevalensi DM tipe 2. Metode ANN merupakan metode berbasis kecerdasan buatan (Artificial Intelligence/AI) dengan menggunakan konsep yang terdiri dari sejumlah besar unit pemrosesan sederhana yang disebut neuron atau node, yang terhubung satu sama lain melalui koneksi bobot menghasilkan. Model ANN untuk prevalensi DM tipe 2 tahun 2019 sampai 2022 adalah ANN dengan 2 hidden layer yang merupakan model yang mampu merepresentasikan prevalensi DM tipe 2 selama periode 2019 sampai 2022. Hasil ini dapat memberikan gambaran bahwa metode-metode berbasis AI dapat digunakan dalam pemodelan dalam bidang kesehatan dan memiliki keunggulan dibandingkan metode statistika pada umumnya yang membutuhkan asumsi-asumsi yang harus dipenuhi. Metode seperti ANN juga dapat dikembangkan selain kasus time series atau prediksi suatu kejadian penyakit, metode ini juga memiliki kemampuan untuk memodelkan dalam kasus klasifikasi untuk menganalisis faktor resiko suatu penyakit pada atau pada kasus spasial untuk mengetahui pola penyebaran penyakit di bidang epidemiolog.

ACCEPTANCE LETTER

Dear Mr./Mrs./Ms. Eva Firdayanti Bisono

It is my great pleasure to inform you that your contribution to JUITA: Jurnal Informatika,

Reference No. : 19564

Title : LENGTH OF STAY (LOS) PREDICTION OF TYPE 2 DIABETES MELLITUS USING CLASSIFICATION AND REGRESSION TREE (CART)

Author(s) : Eva Firdayanti Bisono, Jerhi Wahyu Fernanda, Ratna Frenty Nurkhalim, Krisnita Dwi Jayanti

has been accepted for the publication on Vol. 12 No. 1, Mei 2024.

We would like you to refer to JUITA publications especially your own works in your future research. We believe it will help you promote your article and enhance its visibility.

Best Regards,

Assoc. Prof. Dr. Hindayati Mustafidah

Editor-in-Chief, JUITA

Universitas Muhammadiyah Purwokerto

E-mail: jurnal.juita@ump.ac.id

Purwokerto, December 12, 2023



Assoc. Prof. Dr. Hindayati Mustafidah

LENGTH OF STAY (LOS) PREDICTION OF TYPE 2 DIABETES MELLITUS USING CLASSIFICATION AND REGRESSION TREE (CART)

Eva Firdayanti Bisono¹, Jerhi Wahyu Fernanda², Ratna Frenty Nurkhalim³, Krisnita Dwi Jayanti⁴

^{1,3,4} *Institut Ilmu Kesehatan Bhakti Wiyata Kediri*

² *Institut Agama Islam Negeri Kediri*

eva.firdayanti@iik.ac.id; fernanda.jerhi@iainkediri.ac.id; ratna.nurkhalim@iik.ac.id; krisnita.jayanti@iik.ac.id

Abstract

The prediction of LOS in type 2 patients and the influencing factors can be used as a basis for managing comorbidities and the risk of complications in patients. Predictions can be made using machine learning methods such as Classification and Regression Tree (CART). This study aims to analyse the factors that influence the LOS of type 2 DM patients. The research data was obtained from the Hospital Information System in the period 2019 to 2021 and obtained data for 541 type 2 DM patients. The study variables consisted of the dependent variable, namely LOS of DM patients type 2 and the independent variables consisted of gender, age, complications, comorbidities and urban status of type 2 DM patients. The average LOS of type 2 DM patients was 3.39 days with a median of 3 days. The results of the analysis using CART with 10-fold cross validation concluded that the morbidity variable was the variable that most dominantly influenced the LOS of type 2 DM patients. Accuracy, precision, recall, and F1 scores were respectively 0.704, 0.814, and 0.755.

Keywords: Classification and Regression Tree, Length of Stay, Prediction, Type 2 Diabetes Mellitus

I. INTRODUCTION

In the digitalization of the health services era, providing a good service in a hospital is highly prioritized because it improves the quality of health services. A patient certainly hopes for fast and appropriate services. Shorter waiting times for the patients will increase patient satisfaction. Therefore, skilled professionals are needed. One indicator used for the standard of professionalism of health services and information is indicators in health statistics such as Length of Stay (LOS)[1].

Length of Stay (LOS) is defined as the total number of days a patient has been in the hospital from being registered as an inpatient until completion of care or discharge from the hospital[2]. LOS is an indicator in health services, which can have an impact on hospital financing and patient satisfaction[3]. LOS predictions can also be used as a basis for planning and managing hospital resources to make them more effective and efficient[4]. Analysis of patient LOS factors in hospitals will have an impact on inpatient management more efficiently[5]. In some cases of chronic diseases such as Type 2 Diabetes Mellitus, LOS has a significant influence on the real costs of patient care[6].

Type 2 Diabetes Mellitus is a chronic disease whose prevalence rate always increases every year. Type 2 Diabetes Mellitus in the International Classification of Diseases (ICD) 9 is coded with code E11. **Indonesian Basic Health Research** results in 2018 showed that the prevalence of Diabetes Mellitus was 2.0%. This prevalence value has increased compared to 2013 with a prevalence of 1.5% [7]. This condition will indirectly increase the burden of care costs, especially for BPJS health because this disease needs large costs[8]. Steps that can be taken are to carry out accurate patient identification. To be sure, we must find out some factors that can change LOS with type DM. This can make early warning so that patients get more intensive care so that they can accelerate their recovery and improve other aspects that have an impact on more efficient costs. LOS can also be used to detect the risk of complications and comorbidities in patients with type 2 DM [9].

CART is a machine learning method with supervised learning. CART can be used to predict LOS of Type 2 DM patients by generating a rule based on existing data or information to clarify the relationship between input and output variables [10]. CART also visualizes in the form of a tree diagram that provides an overview of input variables starting with the most important variables used to predict an output. The CART study conducted by Williams, D., et. al used this method for the classification of Parkinson's Disease. Fernanda, J., W. also used CART to analyze risk factors for hypertension[11], [12]. The research results from both studies give a representation that the CART method is very

flexible to use in classification because it is easy to visualize so it can be easily understood to explain the predictor variables. According to Eskandari, M., et, al, we can use the decision tree method to detect what factors can determine LOS in patients suffering from type 2 DM. Therefore, this study aims to determine the early LOS in patients who experience symptoms of type 2 DM according to some of the factors above.[13].

II. METHOD

The research data is secondary data. It obtained from one of the Private Hospital in Kediri. Data imported from the **Hospital Information Management System (HIMS)** starting from 2019 to 2021. This research variables consist of dependent (target) variable namely the Length of Stay (LOS) variable for type 2 DM patients. The independent variable consist of 5 variables namely gender, age, complication status, comorbidities, and patient urban status.

TABLE I
DESCRIPTION OF RESEARCH VARIABLES CATEGORIES

Variables	Categories
Length of Stay (LOS)	<=median
	> median
Gender	Male
	Female
Age	<=45 years
	> 45 years
Complications	Yes
	No
Comorbidities	Yes
	No
Urban Status	Yes
	No

The research data obtained from HIMS is raw data. The data consists of 33 variables in HIMS and needs preprocessing data for carrying data as in Table 1 above. Preprocessing data is very important things because not all imported variables are needed for analysis and there are several variables such as complication status, comorbidity, and urban status are variables from preprocessing another variable. According to Sun, W., et. al. is very necessary to obtain accurate data analysis results[14]. Hassler, A. P., et. al. also stated that data preprocessing is an important stage in the data analysis process in healthcare[15].

The flow of data analysis in this research is divided into two steps, Step 1 is data preprocessing and Step 2 is modelling using the Classification and Regression Tree (CART) method. The preprocessing steps consist of several processes to obtain the variables used in CART modelling. The flow of data analysis can be seen in Fig. 1.

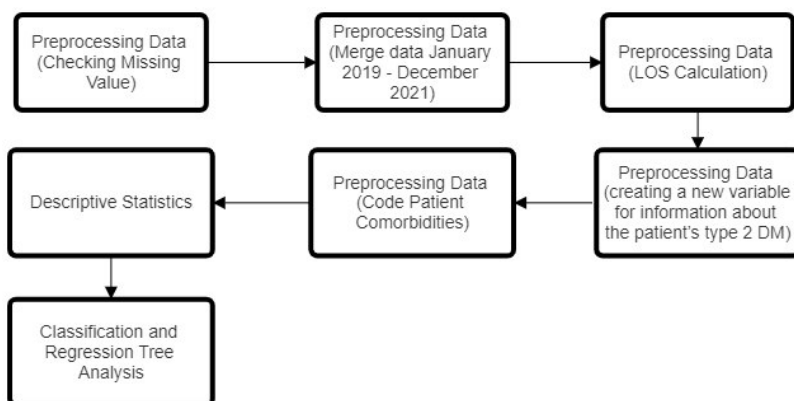


Fig. 1 Flowchart of Data Analysis

A comprehensive explanation of these steps is as follows:

1. Preprocessing Data

Combining data is the first process that must be done. The data originating from HIMS in the form of Excel and separated every month from January 2019 to December 2021 period. The steps in this stage are as follows:

- a. Initial data processing begins by checking missing values. Missing values are checked based on the primary diagnosis variable. At this stage ensure that the patient's primary diagnosis data is filled in and there are no blanks. After ensuring data does not have missing values, the next process is to check the type of variables in the study, especially the variables of patient entry date, patient discharge date, and patient date of birth. There are three variables that are still in the text form so the transformation process to become a date type is carried out. The date variable type for the patient's discharge date and admission date is very necessary for calculating patient LOS.
- b. The next stage is to merge data from January 2019 to December 2021 into one dataset. The data has become a dataset, and the research variables are selected before. The research variables are date admissions, the patient's birth date, gender, primary diagnosis, secondary diagnoses (dtd2 and dtd3), the name of the sub-district, and the name of the patient district. After that, a filter is carried out, especially on dtd1, to filter type 2 DM patients only. Since 2019 to 2021, there were 541 patients suffering from type 2 DM. One of the ways that must be done to find out which patients have this is the diagnosis of E11
- c. The LOS variable is used to calculate the patient's admission and discharge dates.. After that, the LOS variable was coded into two categories, namely \leq median and $>$ median.
- d. The next step is creating a new variable for information about the patient's type 2 DM, whether only type 2 DM without complications or with complications. E11.9 is Type 2 DM without complications, while type 2 DM with codes E11.1 to E11.8 is DM with complications. Codes E11.1 to E11.8 are used to diagnose patients with type 2 DM. Patients who have these codes have complications (1), but if the code is E11.9 then the patient has secondary diagnoses and does not have complications.
- e. To find out which patients have comorbidities by looking at the comorbidity variable and patients who have comorbidities or no comorbidities are coded 0.. The urban variable is a variable obtained from processing the patient's area of origin. Patients come from Kediri district and Pare sub-district, so these patients are categorized as urban, apart from that, patients are categorized as non-urban.

2. Descriptive Statistics

This stage of analysis was carried out to see the description of the variables in the research. Histograms are used to explain LOS data for type 2 DM patients during the period 2019 to 2021. Frequency Distribution Tables are also used to explain the variables of gender, age, complications, morbidity, and urban status

3. Classification and Regression Tree (CART) Analysis

Classification and Regression Tree (CART) is a statistical method used in modelling cases in health, industry, and other fields. CART is a representative of the relationship between the dependent variable (target) and the independent variable which is visualized through a tree diagram, where this tree diagram makes it easy to interpret the results of data analysis [16]. CART consists of nodes or branches and leaves. The top node or root is the variable that has the greatest level of influence. After the node will appear a leaf which represents the class/category of the target variable. In the process of forming this node, a split process is carried out using the Gini Index algorithm. The Gini index algorithm is as follows[17].

$$Gini(D) = 1 - \sum_{i=1}^m p_i^2$$

m is the number of classes. In the CART method, classes are divided into two (binary).

p_i is the probability of sample I to be included in class group m .

In this research, the CART model was validated using the cross-validation (CV) method. CV is functioned to test the model's ability to predict new data that is not visible and is not used in model construction. CV is a method that can be used to obtain optimal machine learning models[18]. This research used the 10-fold cross-validation technique with this following stages below:

- a. Data divided into training data and testing data with proportion number 80% and 20 %
- b. The training data is 80% (432 data), randomized into 10 new datasets that will be used in CART modeling with each dataset also divided into 800% training data and 20% testing data.
- c. CART modelling was carried out on each dataset starting from dataset 1 to 10 and the error level of each model was recorded
- d. Take the average score of 10 CART models from 10 datasets to get a CART model.

4. CART Model Accuracy

After obtaining CART model, prediction is done using the testing data. The prediction results of testing data are presented using the Confusion Matrix Table and based on this table, it is used to measure the goodness of the CART Model [19]. The measures used are accuracy, recall, precision, and F1 Scores [20]. These measurements are also conducted by research [21] to see the performance or accuracy of the machine learning model.

III. RESULT AND DISCUSSION

Type 2 DM patient’s data from 2019 to 2021 is 541 patients. The concentration and distribution of data on the Length of Stay (LOS) of type 2 DM patients is presented in the histogram in Figure 2. The histogram provides information on the LOS of many types 2 DM patients which is 3-4 days. There are also LOS for type 2 DM patients who have a LOS more than 10 days.

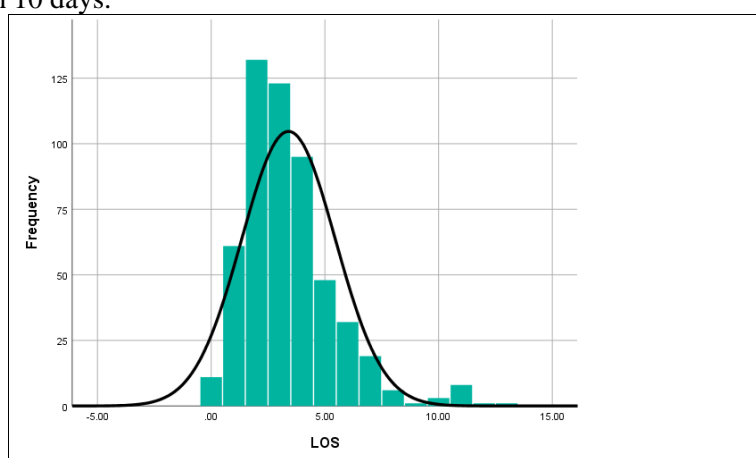


Fig. 2 Patients Type 2 DM LOS Histogram

Descriptive Statistics of LOS for Type 2 DM patients can be seen in Table 2. Patients' LOS average for type 2 DM is 3.39 days with median value of 3 days. The standard deviation of LOS Type 2 DM patients is 2.06. The median value will be used as the basis for categorizing LOS patients which used in the analysis using CART method. In the analysis using CART, LOS data will be categorized into ≤ 3 days and >3 days.

**TABLE II
DESCRIPTIVE STATISTICS PATIENTS LOS DM TYPE 2**

Variable	Mean	Median	Standard Deviation
Length of Stay (LOS)	3.39	3	2.06

Respondent characteristics based on independent variables such as gender, age, comorbidities, complications, and urban area are presented in Table 3 which explains the frequency distribution of each variable. The majority case of Type 2 DM patients in this research where female with percentage is 62.9% and 37.2% were male.

**TABLE III
FREQUENCY DISTRIBUTION OF PREDICTOR VARIABLES**

Variable	Category	Frequency	Percentage
Gender	Female	340	62.8
	Male	201	37.2
Age	<=45	49	9.1
	>45	492	90.9
Comorbidities	Yes	121	22.4
	No	420	77.6
Complications	Yes	211	39
	No	330	61
Urban Status	Yes	105	19.4
	No	436	80.6

The majority, of type 2 DM patients are over 45 years old with a 90.05 percent. In this research, were also found patients suffering from type 2 DM aged <= 45 years with a percentage of 9.1%. There were 121 patients (22.4%) who had comorbidities. Of 541 type 2 DM patients, 211 patients (39%) have complications. Patients are categorized into urban and non-urban based on the area where they live. There were 105 patients with type 2 DM who came from urban areas (19.4%) and 436 patients who did not come from urban areas (80.6%).

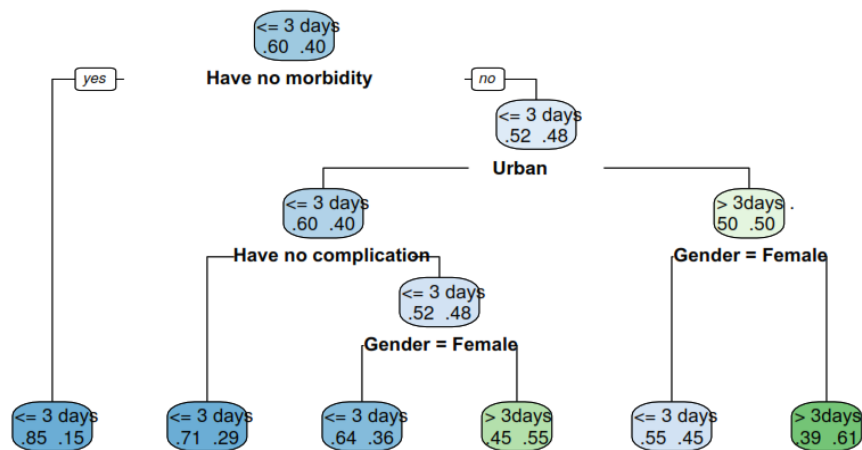


Fig. 3 Classification and Regression Tree (CART) Structure of Type 2 DM Patients

Fig. 3 is the CART Structure produced to predict LOS for type 2 DM patients. The morbidity variable is the first node in the CART structure. This condition indicates that the patient's morbidity variable is the most influential variable on the LOS of Type 2 DM patients. Patients who do not have comorbidities will have a great chance of having a LOS of less than or equal to 3 days. If the patient has comorbidities, then to predict LOS for type 2 DM patients, variables such as urban area, complications, and gender need to be taken into account.

Patients who do not have comorbidities, come from urban areas, and do not have complications experienced, have a high chance of having a LOS <= 3 days. Meanwhile, patients who do not have comorbidities, are not from urban areas, and are female will have a chance of having a LOS of <= 3 days, and if they are male, they will have a chance of having a LOS > 3 days.

**TABLE III
CONFUSION MATRIX PREDICTING TESTING DATA**

Prediction	Actual	
	<=3 days	> 3 days
< 3 days	57	24
> 3 days	13	14

The confusion matrix table in table III explains that there are 24 data which in the real data are LOS ≤ 3 days, predicted using the CART method to be > 3 days. Meanwhile, there were 13 actual data on patient LOS > 3 days which were predicted by the CART model to be ≤ 3 days. The confusion matrix table is the basis for calculating accuracy and F1 score values from the CART model.

TABLE IV
CART MODEL PERFORMANCE

Parameter	F1 Score
Accuracy	0.657
Precision	0.704
Recall	0.814
F1 Score	0.755

The accuracy value in this study was 0.6574 or 65.74% due to the complex characteristics of the data with various variables. Meanwhile, the precision, recall and F1 Score values are 0.704, 0.814 and 0.755.

IV. CONCLUSION

The results of the analysis using the CART method provide the conclusion that the variables morbidity, urbanity, complications, and gender are variables that can be used to predict LOS for type 2 DM patients. The morbidity variable is the variable with the highest level of importance compared to other variables. The CART model has an accuracy of 0.6574 (65.74%), and an F1 Score of 0.76.

ACKNOWLEDGEMENT

We thank to Ministry of Research and Higher Education, RS HVA Kediri, IIK Bhakti Wiyata Kediri, and IAIN Kediri who have help this research and funders for the research carried out or to people who are truly instrumental in carrying out the research.

REFERENCES

- [1] R. Rosita and A. R. Tanastasya, "Penetapan mutu rumah sakit berdasarkan indikator rawat inap," *J. Kesehat. Kusuma Husada*, pp. 166–178, 2019.
- [2] M. J. Bowie, *Essentials of health information management: Principles and practices*. Cengage Learning, 2022.
- [3] J. Chrusciel, F. Girardon, L. Roquette, D. Laplanche, A. Duclos, and S. Sanchez, "The prediction of hospital length of stay using unstructured data," *BMC Med. Inform. Decis. Mak.*, vol. 21, no. 1, p. 351, 2021.
- [4] K. Stone, R. Zwigelaar, P. Jones, and N. Mac Parthaláin, "A systematic review of the prediction of hospital length of stay: Towards a unified framework," *PLOS Digit. Heal.*, vol. 1, no. 4, p. e0000017, 2022.
- [5] H. Baek, M. Cho, S. Kim, H. Hwang, M. Song, and S. Yoo, "Analysis of length of hospital stay using electronic health records: A statistical and data mining approach," *PLoS One*, vol. 13, no. 4, p. e0195901, 2018.
- [6] I. Oktadiana, "Perbandingan Biaya Riil Pada Pasien Diabetes Mellitus Tipe 2 Dengan Tarif INA-CBG'S Di Rumah Sakit Umum Daerah," *J. Farm. Tinctura*, vol. 2, no. 2, pp. 42–51, 2021.
- [7] I. Irwansyah and I. S. Kasim, "Identifikasi keterkaitan lifestyle dengan risiko diabetes melitus," *J. Ilm. Kesehat. Sandi Husada*, vol. 10, no. 1, pp. 62–69, 2021.

- [8] A. E. Pratiwi and H. Sukmawati, "ANALISIS BIAYA RATA-RATA PASIEN RAWAT INAP DENGAN PENYAKIT DIABETES MELLITUS TYPE II (STUDI DI JEMBRANA DAN GIANYAR)," *WICAKSANA J. Lingkungan. dan Pembang.*, vol. 3, no. 2, pp. 21–29, 2019.
- [9] B. Katipoglu, M. I. Naharci, and E. S. Yurdakul, "Risk factors predicting hospital length of stay in older patients with type 2 diabetes with Covid-19," *J. Diabetes Metab. Disord.*, vol. 21, no. 2, pp. 1443–1449, 2022.
- [10] M. Ozcan and S. Peker, "A classification and regression tree algorithm for heart disease modeling and prediction," *Healthc. Anal.*, vol. 3, p. 100130, 2023.
- [11] D. Williams *et al.*, "Visualization of Decision Tree State for the Classification of Parkinson's Disease," *J. Biomed. Eng. Med. Imaging*, vol. 3, no. 3, pp. 25–41, 2016.
- [12] J. W. Fernanda, G. Anuraga, and M. A. Fahmi, "Risk factor analysis of hypertension with logistic regression and Classification and Regression Tree (CART)," in *Journal of Physics: Conference Series*, 2019, vol. 1217, no. 1, p. 12109.
- [13] M. Eskandari, A. H. Alizadeh Bahmani, H. A. Mardani-Fard, I. Karimzadeh, N. Omidifar, and P. Peymani, "Evaluation of factors that influenced the length of hospital stay using data mining techniques," *BMC Med. Inform. Decis. Mak.*, vol. 22, no. 1, pp. 1–11, 2022.
- [14] W. Sun, Z. Cai, Y. Li, F. Liu, S. Fang, and G. Wang, "Data processing and text mining technologies on electronic medical records: a review," *J. Healthc. Eng.*, vol. 2018, 2018.
- [15] A. P. Hassler, E. Menasalvas, F. J. García-García, L. Rodríguez-Mañas, and A. Holzinger, "Importance of medical data preprocessing in predictive modeling and risk factor discovery for the frailty syndrome," *BMC Med. Inform. Decis. Mak.*, vol. 19, pp. 1–17, 2019.
- [16] C. Machuca, M. V Vettore, M. Krasuska, S. R. Baker, and P. G. Robinson, "Using classification and regression tree modelling to investigate response shift patterns in dentine hypersensitivity," *BMC Med. Res. Methodol.*, vol. 17, pp. 1–11, 2017.
- [17] T. Daniya, M. Geetha, and K. S. Kumar, "Classification and regression trees with gini index," *Adv. Math. Sci. J.*, vol. 9, no. 10, pp. 8237–8247, 2020.
- [18] I. K. Nti, O. Nyarko-Boateng, and J. Aning, "Performance of machine learning algorithms with different K values in K-fold cross-validation," *J. Inf. Technol. Comput. Sci.*, vol. 6, pp. 61–71, 2021.
- [19] E. S. Kresnawati, Y. Resti, B. Suprihatin, M. R. Kurniawan, and W. A. Amanda, "Coronary Artery Disease Prediction Using Decision Trees and Multinomial Na⁺ ve Bayes with k-Fold Cross Validation," *Inomatika*, vol. 3, no. 2, pp. 172–187, 2021.
- [20] R. M. AlZoman and M. J. F. Alenazi, "A comparative study of traffic classification techniques for smart city networks," *Sensors*, vol. 21, no. 14, p. 4677, 2021.
- [21] A. PINAR, C. Çolak, and E. Gültürk, "Evaluation of Performance Metrics in Heart Disease by Machine Learning Techniques," *J. Cogn. Syst.*, vol. 8, no. 1, pp. 11–15.

PREDIKSI PREVALENSI DIABETES TIPE 2 MENGUNAKAN ARTIFICIAL NEURAL NETWORK

PENELITIAN DOSEN PEMULA

RS HVA TOELOENGREDJO PARE KEDIRI

Eva Firdayanti Bisono, Krisnita Dwi Jayanti, Ratna Frenty Nurkhalim

IIK BHAKTI WIYATA KEDIRI

2023



Detail Penelitian



- Judul Penelitian : Prediksi Prevalensi Diabetes Tipe 2 Menggunakan Artificial Neural Network
- Lokasi Penelitian : RS HVA Toeloengredjo Pare
- Nama Pengusul :
 - Eva Firdayanti Bisono, M.Kom (0711069201)
 - Ratna Frenty Nurkhalim, S.KM., M.P.H (0729079002)
 - Krisnita Dwi Jayanti, S.KM., M.Epid (0718108802)
- Tahun Pelaksanaan : 2023
- Institusi : IIK Bhakti Wiyata Kediri

- **Latar Belakang**

- Peningkatan jumlah penderita DM Tipe 2 sebesar 2% di tahun 2021 dibandingkan tahun 2020

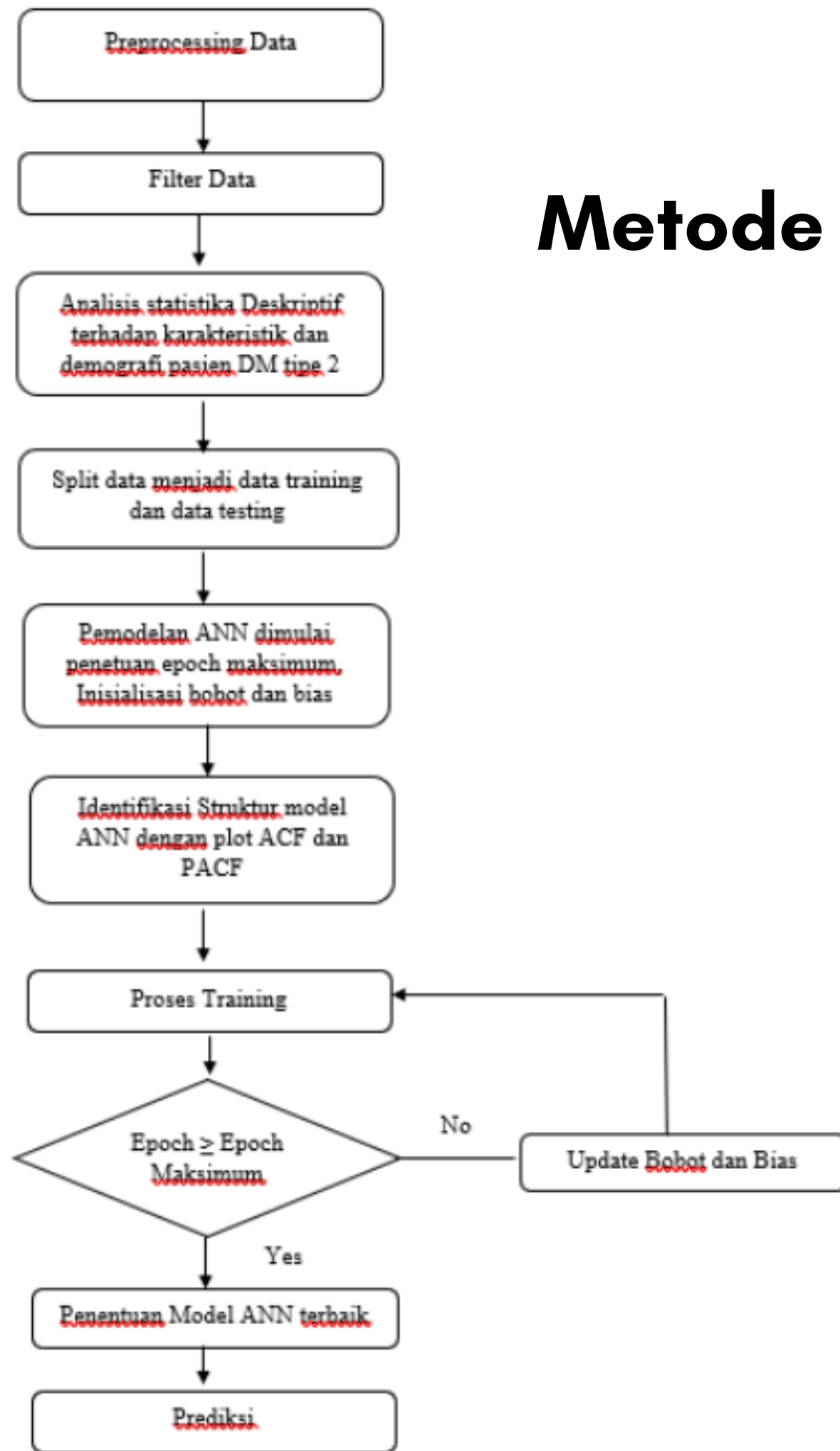
- **Tujuan**

- Membantu sebagai early warning system yang memberikan respon kepada epidemiologis yang cepat sehingga mampu mengurangi morbidity dan mortality

- **Kebaruan**

- Metode forecasting di bidang epidemiologi menggunakan metode ANN di Indonesia masih sangat terbatas
- Metode yang banyak digunakan adalah ARIMA. Namun ARIMA membutuhkan asumsi yang harus dipenuhi.

Metode



• Hasil Penelitian

Tabel 8. Prediksi prevalensi pada data testing dengan variasi jumlah neuron

Bulan	Aktual Prevalensi	Jumlah Neuron pada hidden layer								
		2	3	4	5	6	7	8	9	10
Juli	0.9901	0.7856	0.7888	0.8054	0.7807	0.7815	0.7887	0.7855	0.7948	0.7959
Agustus	0.0724	0.5693	0.5791	0.5890	0.5821	0.5810	0.5801	0.5820	0.5842	0.5873
September	0.4698	0.8223	0.8270	0.8549	0.8136	0.8146	0.8266	0.8225	0.8383	0.8392
Oktober	0.2708	0.6716	0.6864	0.7103	0.6924	0.6901	0.6928	0.6912	0.6973	0.7020
November	0.2751	0.8455	0.8508	0.8863	0.8340	0.8350	0.8500	0.8453	0.8665	0.8666
Desember	0.0000	0.7346	0.7490	0.7899	0.7594	0.7571	0.7626	0.7585	0.7697	0.7732

Tabel 9. Nilai RMSE model ANN dengan neuron yang berbeda pada hidden layer

RMSE	Jumlah Neuron pada hidden layer								
	2	3	4	5	6	7	8	9	10
RMSE	0.4912	0.6414	0.8057	0.9270	1.0691	1.2165	1.3576	1.5086	1.6533

Tabel 9. Prediksi prevalensi menggunakan 2 neuron pada hidden layer.

Bulan	Prediksi
Januari 2023	0.776
Februari 2023	0.576
Maret 2023	0.806
April 2023	0.681
Mei 2023	0.825
Juni 2023	0.745

Luaran

- **Jurnal Terakreditasi Sinta 2**
- **Status : Accepted**
- **JUITA : Jurnal Informatika**

- **Buku Ajar**
- **Status : Penerbit**

Kesimpulan

- Hasil prediksi cenderung stagnan dan tidak terdapat pola trend menurun.
- Metode baik dalam memprediksi prevalensi diabetes tipe 2 terbaik pada neuron 2, hal ini dapat dilihat dari nilai RMSE berada < 50
- Prevalensi Diabetes Tipe 2 terbesar yaitu berada pada bulan Mei 2023

Foto Aktivitas Penelitian



SURAT PERNYATAAN TANGGUNG JAWAB BELANJA

Yang bertanda tangan di bawah ini :

Nama : EVA FIRDAYANTI BISONO S.Kom, M.Kom

Alamat : TAMAN PINANG INDAH G-VII/12A

berdasarkan Surat Keputusan Nomor 199/SPK/D.D4/PPK.01/APTV/VI/2-23 dan Perjanjian / Kontrak Nomor 011/SP2H/PPKM-PTV/LL7/2023, 596/R/PN/VI/2023 mendapatkan Anggaran Penelitian Prediksi Prevalensi Diabetes Tipe 2 menggunakan Artificial Neural Network Sebesar 11,860,000 Dengan ini menyatakan bahwa :

1. Biaya kegiatan Penelitian di bawah ini meliputi :

No	Uraian	Jumlah
01	Bahan ATK	360,000
02	Pengumpulan Data Transport Peneliti, HR Pembantu Peneliti dan Pembantu Lapangan	7,460,000
03	Analisis Data(Termasuk Sewa Peralatan HR Pengolah data dan pembuatan pemetaan	1,540,000
04	Pelaporan, Luaran Wajib dan Luaran Tambahan Jurnal dan Buku	2,350,000
05	Lain-lain Etik Penelitian	150,000
	Jumlah	11,860,000

2. Jumlah uang tersebut pada angka 1, benar-benar dikeluarkan untuk pelaksanaan kegiatan Penelitian dimaksud.

Demikian surat pernyataan ini dibuat dengan sebenarnya.

Kediri, 14-12-2023

Ketua,



(EVA FIRDAYANTI BISONO S.Kom, M.Kom)

NIP/NIK 3515085106920002