



KARYA TULIS ILMIAH

**PREDIKSI PENYAKIT DIABETES MELLITUS DENGAN *MACHINE*
LEARNING DI RSUD dr. SLAMET GARUT**

MUHAMMAD GIBBRAN AKBARY

P2.06.37.0.22.060

PROGRAM STUDI DIPLOMA TIGA
REKAM MEDIS DAN INFORMASI KESEHATAN TASIKMALAYA

JURUSAN REKAM MEDIS DAN INFORMASI KESEHATAN
POLITEKNIK KESEHATAN TASIKMALAYA
KEMENTERIAN KESEHATAN REPUBLIK INDONESIA
TAHUN 2024

UNGKAPAN TERIMA KASIH

Dengan ucapan syukur Alhamdulillah kehadiran Allah SWT atas rahmat dan hidayah Nya, yang telah memberikan petunjuk dan kemudahan bagi kami sehingga kami dapat menyelesaikan proposal dengan judul Prediksi Penyakit Diabetes Dengan *Machine Learning* di RSUD dr. Slamet Garut. Penyusunan proposal ini terwujud atas bimbingan, pengarahan dan bantuan dari berbagai pihak. Oleh karena itu pada kesempatan ini penulis menyampaikan penghargaan dan terima kasih kepada:

1. Direktur Poltekkes Tasikmalaya, Ibu Dr. Dini Mariani, S.Kep., Ners M.Kep;
2. Ketua Jurusan Program Studi D III Rekam Medis dan Informasi Kesehatan, Bapak Andi Suhenda SKM.MPH;
3. Direktur Rumah Sakit Umum Daerah (RSUD) dr. Slamet Garut, dr. H. Husodo Dewo Adi, SP.OT-Spine;
4. Petugas Rekam Medis Rumah Sakit Umum Daerah (RSUD) dr. Slamet Garut;
5. Dosen Pembimbing, Ibu Diana Barsasella, MKM, Ph.D;
6. Seluruh Dosen dan Staff Jurusan Rekam Medis dan Informasi Kesehatan Poltekkes Kemenkes Tasikmalaya;
7. Orang tua dan keluarga yang telah mendo'akan serta memberi dukungan dalam pembuatan proposal penelitian ini;
8. Teman-teman yang selalu memberi dukungan dan semangat dalam penyelesaian proposal ini.

Penulis menyadari sepenuhnya bahwa penyusunan proposal ini masih banyak kekurangan dan jauh dari sempurna, oleh karena itu penulis dengan tulus meminta kritik dan saran yang membangun dari pembaca sehingga digunakan untuk perbaikan lebih lanjut. Penulis berharap semoga laporan ini bermanfaat bagi penulis khususnya serta bagi pembaca umumnya.

Tasikmalaya, Januari 2025

Penulis

Kementerian Kesehatan Republik Indonesia
Poltekkes Kemenkes Tasikmalaya
Jurusan Rekam Medis dan Informasi Kesehatan
Program Studi Diploma III Rekam Medis dan Informasi Kesehatan
Tasikmalaya
2025
MUHAMMAD GIBBRAN AKBARY

**PREDIKSI PENYAKIT DIABETES MELITUS DENGAN *MACHINE LEARNING* DI RSUD dr.
SLAMET GARUT**

98 Hal, 5 Bab, 9 Tabel, 8 Gambar, 23 Lampiran

ABSTRAK

Latar Belakang: Diabetes mellitus adalah suatu kondisi yang ditandai dengan tingginya kadar gula (glukosa) dalam darah secara terus-menerus. Prediksi penyakit diabetes mellitus dapat membantu dalam pendeteksian dini kasus, untuk segera diupayakan intervensi yang tepat dan cepat, sehingga dapat meningkatkan kualitas hidup seseorang dan menghindari komplikasi yang dapat terjadi.

Metode: Penelitian ini menggunakan analisis kuantitatif dengan pendekatan metode *data mining* yaitu proses penggalian sebuah informasi dan pengetahuan baru dari data-data yang tersimpan pada *dataset*. Variabel yang digunakan pada *dataset* mencakup data terkait rekam medis dan hasil pemeriksaan laboratorium, pada kunjungan pasien periode Juni – Desember 2024 di RSUD dr. Slamet Garut

Hasil Penelitian: Hasil menggunakan widget Test and Score dengan menggunakan Test on test data atau menggunakan file test yang berbeda dengan file yang digunakan dalam training untuk dijadikan data test. Score hasil menunjukkan bahwa *Naive Bayes* unggul secara keseluruhan. *Naive Bayes* memiliki AUC tinggi (0.952), F1 score dan precision yang lebih baik, serta MCC sebesar 0.542, menandakan prediksi yang lebih akurat dan seimbang. Akurasi Klasifikasi (CA) *Naive Bayes* memang sedikit lebih rendah (77.8%), namun metrik lainnya menunjukkan performa yang jauh lebih stabil dan dapat diandalkan andal dengan data yang tidak seimbang.

Kesimpulan: Hasil penelitian ini menunjukkan bahwa hasil dari machine learning, khususnya model *Naive Bayes*, dapat memberikan kontribusi signifikan dalam mendukung dalam melakukan prediksi apakah pasien terdiagnosis diabetes berdasarkan pemeriksaan laboratoriumnya. Prediksi diabetes dengan model prediksi berdasarkan probabilitas (*Naive Bayes*) lebih efektif dalam pelaksanaannya karena hasil laboratorium dapat diklasifikasikan berdasarkan kemungkinan apakah hasil yang didapat normal atau tidak.

Kata Kunci: diabetes mellitus, *machine learning*, *data mining*, *orange data mining*, prediksi penyakit

Ministry of Health of the Republic of Indonesia
Poltekkes Kemenkes Tasikmalaya
Department of Medical Records and Health Informations
Diploma III Study Program in Medical Records and Health Information
Tasikmalaya
2025
MUHAMMAD GIBBRAN AKBARY

**DIABETES MELLITUS DISEASE PREDICTION WITH MACHINE LEARNING AT dr. SLAMET
GARUT HOSPITAL**

98 Page, 5 Chapter, 9 Table, 8 Figure, 23 Attachment

ABSTRACT

Background: Diabetes mellitus is a condition characterized by persistently high levels of sugar (glucose) in the blood. Prediction of diabetes mellitus disease can help in early detection of cases, to immediately attempt appropriate and fast intervention, so as to improve a person's quality of life and avoid complications that may occur.

Method: This study uses quantitative analysis with a data mining method approach, namely the process of extracting new information and knowledge from data stored in the dataset. The variables used in the dataset include data related to medical records, laboratory test results, and patient pharmacy data at patient visits for the period June - December 2024 at RSUD dr. Slamet

Research Results: The results of using the Test and Score widget using Test on test data or using a different test file from the file used in training to be used as test data. Naive Bayes has a high AUC (0.952), better F1 score and precision, and an MCC of 0.542, indicating a more accurate and balanced prediction. Naive Bayes Classification Accuracy (CA) is indeed slightly lower (77.8%), but other metrics show much more stable and reliable performance with imbalanced data.

Conclusion: The results of this study indicate that the results of machine learning, especially the Naïve Bayes model, can provide a significant contribution in supporting the prediction of whether a patient is diagnosed with diabetes based on their laboratory examination. Diabetes prediction with a probability-based prediction model (Naïve Bayes) is more effective in its implementation because laboratory results can be classified based on the possibility of whether the results obtained are normal or not.

Keywords: diabetes mellitus, machine learning, data mining, orange data mining, disease prediction

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN	iv
UNGKAPAN TERIMAKASIH	vi
MOTTO	vii
PERSEMBAHAN.....	viii
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiv
DAFTAR TABEL.....	xv
DAFTAR LAMPIRAN.....	xvi
BAB I PENDAHULUAN.....	1
A. Latar Belakang	1
B. Tujuan.....	4
1. Tujuan Umum.....	4
2. Tujuan Khusus	5
C. Manfaat.....	5
1. Praktisi	5
2. Teoritis.....	5
D. Keaslian Penelitian	6
BAB II TINJAUAN PUSTAKA	8
A. Telaah Teori	8
1. Penyakit Tidak Menular	8
2. Diabetes Mellitus.....	8
3. <i>Orange Data Mining</i>	9
4. <i>Machine Learning</i>	11
5. <i>Data Mining</i>	11
6. <i>Data Preprocessing</i>	12
7. Model Algoritma: <i>Naïve Bayes</i>	12
8. Model Algoritma: <i>K-Nearest Neighbor</i>	14
9. Evaluasi Model	16
B. Kerangka Teori.....	20
C. Kerangka Konsep	20
BAB III METODE PENELITIAN	21

A.	Jenis dan Desain Penelitian	21
B.	Tempat dan Waktu Penelitian	21
C.	Populasi Sampel dan Teknik Sampel	21
1.	Populasi	21
2.	Sampel	21
3.	Teknik Sampling.....	22
D.	Variabel Penelitian	22
1.	Variabel Independen/Bebas	22
2.	Variabel Dependen/Terikat.....	23
E.	Definisi Operasional.....	23
F.	Instrumen dan Cara Pengumpulan Data	26
1.	Instrumen Penelitian	26
2.	Cara Pengumpulan Data	27
G.	Pengolahan Data.....	27
1.	Pengumpulan Data.....	27
2.	Data <i>Preprocessing</i>	27
3.	Pemilihan Metode.....	28
4.	Pembagian data.....	28
5.	Model <i>Training</i>	28
6.	Evaluasi Model	29
7.	Evaluasi Hasil	29
H.	Rencana Analisis Data	29
I.	Etika Penelitian	29
1.	<i>Research permit</i>	29
2.	<i>Ethical Clearance</i>	30
3.	<i>Confidentiality</i>	30
4.	<i>Benefit</i>	30
J.	Jalannya Penelitian.....	30
1.	Tahap Persiapan.....	30
2.	Tahap Pelaksanaan	30
3.	Tahap Lanjutan	30
K.	Jadwal Penelitian.....	30
	BAB IV HASIL DAN PEMBAHASAN	32
A.	Hasil Kegiatan	32
1.	Gambaran Umum RSUD dr. Slamet Garut	32

2.	Penjelasan Data.....	33
3.	Pengolahan Data / <i>Data Processing</i>	34
4.	<i>Machine Learning</i>	36
5.	Visualisasi dan Evaluasi Hasil.....	40
B.	Pembahasan.....	43
1.	Hasil Evaluasi Model Algoritma	43
2.	Faktor – Faktor yang Menyebabkan Pasien Dapat Terdiagnosis Diabetes	44
3.	Pendeteksian Dini Diabetes	45
4.	Model Algoritma yang Lebih Baik.....	46
5.	Efektivitas Model Algoritma <i>Naïve Bayes</i>	47
6.	Keterbatasan dalam Pengolahan dan Proses Data	48
BAB V KESIMPULAN DAN SARAN		50
A.	Kesimpulan.....	50
B.	Saran.....	51
1.	Bagi Pihak Rumah Sakit.....	51
2.	Bagi Peneliti Selanjutnya.....	52
DAFTAR PUSTAKA		
DAFTAR RIWAYAT HIDUP		
LAMPIRAN		

DAFTAR GAMBAR

Gambar 2. 1 Rumus Teorema Bayes	13
Gambar 2. 2 Rumus Jarak Euclidean	15
Gambar 2. 3 Kerangka Teori	20
Gambar 2. 4 Kerangka Konsep	20
Gambar 4. 1 Alur Proses <i>Cleaning Data</i>	35
Gambar 4. 2 <i>Workflow Orange</i>	38
Gambar 4. 3 Scatter Plot antara Glukosa Darah Sewaktu dengan Hasil Prediksi	42
Gambar 4. 4 Scatter Plot antara Usia dengan Hasil Prediksi	42

DAFTAR TABEL

Tabel 1. 1 Keaslian Penelitian	6
Tabel 3. 1 Definisi Operasional	23
Tabel 3. 2 Jadwal Penelitian	30
Tabel 4. 1 Tabel Karakteristik	36
Tabel 4. 2 Pembagian <i>Role</i> Variabel.....	37
Tabel 4. 3 <i>Rank Features ReliefF</i>	38
Tabel 4. 4 <i>Matrix Test and Score</i>	39
Tabel 4. 5 <i>Confusion Matrix Naïve Bayes</i>	40
Tabel 4. 6 <i>Confusion Matrix KNN</i>	40

DAFTAR LAMPIRAN

- Lampiran 1 Lembar Riwayat Hidup
- Lampiran 2 Lembar Persetujuan Judul Karya Tulis
- Lampiran 3 Lembar Persetujuan Judul Tugas Akhir
- Lampiran 4 Lembar Bimbingan Proposal KTI
- Lampiran 5 Surat Pengajuan Studi Pendahuluan 1
- Lampiran 6 Surat Pengajuan Studi Pendahuluan 2
- Lampiran 7 Surat Balasan Bakesbangpol 1
- Lampiran 8 Surat Balasan Bakesbangpol 2
- Lampiran 9 Surat Balasan Dinas Kesehatan
- Lampiran 10 Surat Rekomendasi Ujian Proposal KTI
- Lampiran 11 Surat Pengajuan Kaji Etik
- Lampiran 12 Kaji Etik Penelitian
- Lampiran 13 Surat Izin Penelitian 1
- Lampiran 14 Surat Izin Penelitian 2
- Lampiran 15 Surat Balasan Izin Penelitian Kesbangpol
- Lampiran 16 Surat Balasan Izin Penelitian Kesbangpol
- Lampiran 17 Surat Balasan Izin Penelitian Dinas Kesehatan
- Lampiran 18 Data Kunjungan Sebelum Data *Cleaning*
- Lampiran 19 Data Pemeriksaan Lab Sebelum Cleaning
- Lampiran 20 Data *Training*
- Lampiran 21 Data *Testing*
- Lampiran 22 Lembar Bimbingan Hasil Karya Tulis
- Lampiran 23 Surat Rekomendasi Sidang Hasil Karya Tulis