

**IMPLEMENTASI ARTIFICIAL INTELLIGENCE
DALAM PREDIKSI PENYAKIT DIABETES
MENGUNAKAN ALGORITMA BINARY
CLASSIFICATION**

SKRIPSI



**Oleh:
Hanna Yunita
210210062**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2025**

**IMPLEMENTASI ARTIFICIAL INTELLIGENCE
DALAM PREDIKSI PENYAKIT DIABETES
MENGUNAKAN ALGORITMA BINARY
CLASSIFICATION**

SKRIPSI

**Untuk memenuhi salah satu syarat
memperoleh gelar Sarjana**



**Oleh:
Hanna Yunita
210210062**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2025**

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini penulis:

Nama : Hanna Yunita
NPM : 210210062
Fakultas : Teknik dan Komputer
Program Studi : Teknik Informatika

Menyatakan bahwa "Skripsi" yang penulis buat dengan judul:

IMPLEMENTASI ARTIFICIAL INTELLIGENCE DALAM PREDIKSI PENYAKIT DIABETES MENGGUNAKAN ALGORITMA BINARY CLASSIFICATION

Adalah hasil karya sendiri dan bukan "duplikasi" dari karya orang lain. Sepengetahuan penulis, didalam naskah skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip didalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata di dalam naskah skripsi ini dapat dibuktikan terdapat unsur-unsur PLAGIASI. Penulis bersedia naskah skripsi ini digugurkan dan judul penelitian yang penulis peroleh dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku.

Batam, 23 Juli 2025


Hanna Yunita
210210062

**IMPLEMENTASI ARTIFICIAL INTELLIGENCE
DALAM PREDIKSI PENYAKIT DIABETES
MENGUNAKAN ALGORITMA BINARY
CLASSIFICATION**

SKRIPSI

Untuk memenuhi salah satu syarat
memperoleh gelar Sarjana

Oleh:
Hanna Yunita
210210062

Telah disetujui oleh Pembimbing pada tanggal
seperti tertera di bawah ini

Batam, 22 Juli 2025



Elbert Hutabri, S.Kom., M.Kom.
Pembimbing

ABSTRAK

Penyakit diabetes melitus merupakan salah satu penyakit tidak menular yang terus meningkat dan menjadi ancaman serius bagi kesehatan masyarakat, sehingga deteksi dini terhadap risiko diabetes sangat penting untuk mencegah komplikasi yang lebih berat. Penelitian ini bertujuan untuk mengembangkan sistem prediksi risiko diabetes secara otomatis menggunakan pendekatan kecerdasan buatan (*Artificial Intelligence*) dengan metode *binary classification* yang diimplementasikan melalui fitur *AutoAI* dalam IBM Cloud Pak for Data. Dataset yang digunakan bersumber dari Kaggle dan terdiri dari 724 data pasien setelah melalui proses pembersihan. Terdapat delapan parameter input utama, yaitu: *Pregnancies*, *Glucose*, *Blood Pressure*, *Skin Thickness*, *Insulin*, *BMI*, *Diabetes Pedigree Function*, dan *Age*. Data dibagi secara manual menjadi 652 data untuk pelatihan dan 72 data untuk pengujian. *AutoAI* secara otomatis melakukan pemilihan algoritma, rekayasa fitur, serta evaluasi model untuk memperoleh hasil klasifikasi terbaik. Hasil penelitian menunjukkan bahwa model yang dibangun mampu mengklasifikasikan data dengan akurasi sebesar 81,8%, *precision* 66,7%, dan *recall* 90,9%, yang menunjukkan bahwa sistem memiliki kinerja yang cukup baik dalam mengenali risiko diabetes. Model ini juga mampu mempelajari pola-pola kompleks dari data *input* untuk memberikan hasil prediksi yang relevan. Dengan demikian, sistem yang dikembangkan dapat dimanfaatkan sebagai alat bantu deteksi dini risiko diabetes secara efisien dan akurat.

Kata kunci: *Artificial Intelligence*, *AutoAI*, *Binary Classification*, IBM Cloud Pak for Data, Prediksi Diabetes

ABSTRACT

Diabetes mellitus is one of the non-communicable diseases that continues to increase and poses a serious threat to public health, making early detection of diabetes risk is essential to prevent more severe complications. This study aims to develop an automated system for predicting diabetes risk using Artificial Intelligence (AI) with a binary classification approach implemented through the AutoAI feature in IBM Cloud Pak for Data. The dataset used was obtained from Kaggle and consisted of 724 patient records after the data cleaning process. Eight main input parameters were used: Pregnancies, Glucose, Blood Pressure, Skin Thickness, Insulin, BMI, Diabetes Pedigree Function, and Age. The data was manually split into 652 training records and 72 testing records. AutoAI automatically handled algorithm selection, feature engineering, and model evaluation to obtain the best classification results. The results showed that the developed model was able to classify data with an accuracy of 81.8%, precision of 66.7%, and recall of 90.9%, indicating that the system performed well in recognizing diabetes risk. The model was also capable of learning complex patterns from the input data to provide relevant prediction results. Therefore, the developed system can be used as an efficient and accurate tool for early diabetes risk detection.

Keywords: Artificial Intelligence, AutoAI, Binary Classification, IBM Cloud Pak for Data, Diabetes Prediction

KATA PENGANTAR

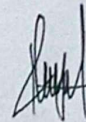
Puji dan syukur kepada Tuhan Yang Maha Esa yang telah melimpahkan segala berkat dan karuniaNya, sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada Program Studi Teknik Informatika Universitas Putera Batam.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Karena itu, kritik dan saran akan senantiasa penulis terima dengan senang hati. Dengan segala keterbatasan, penulis menyadari pula bahwa skripsi ini takkan terwujud tanpa bantuan, bimbingan, dan dorongan dari berbagai pihak. Untuk itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada:

1. Rektor Universitas Putera Batam Ibu Dr. Nur Elfi Husda, S.Kom., M.SI.
2. Dekan Fakultas Teknik dan Komputer Bapak Welly Sugianto, S.T., M.M., Ph.D.
3. Ketua Program Studi Teknik Informatika Bapak Andi Maslan, S.T., M.SI., Ph.D.
4. Bapak Ellbert Hutabri, S.Kom., M.Kom. selaku pembimbing Skripsi pada Program Studi Teknik Informatika Universitas Putera Batam.
5. Dosen dan Staff Universitas Putera Batam.
6. Orangtua dan keluarga yang selalu mendukung dan mendoakan.

Semoga Tuhan Yang Maha Esa membalas kebaikan dan senantiasa menyertai dengan kasih dan hikmat-Nya, Amin.

Batam, 23 Juli 2025



Hanna Yunita

DAFTAR ISI

	Halaman
HALAMAN SAMPUL	i
HALAMAN JUDUL	ii
SURAT PERNYATAAN ORISINALITAS	iii
HALAMAN PENGESAHAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
DAFTAR RUMUS	xii
BAB I PENDAHULUAN	1
1.1. Latar Belakang.....	1
1.2. Identifikasi Masalah	3
1.3. Batasan Masalah	3
1.4. Rumusan Masalah	4
1.5. Tujuan Penelitian.....	4
1.6. Manfaat Penelitian.....	5
1.6.1. Manfaat Teoritis	5
1.6.2. Manfaat Praktis.....	5
BAB II TINJAUAN PUSTAKA	7
2.1. Teori Dasar	7
2.1.1. <i>Artificial Intelligence</i>	7
2.1.2. Pembelajaran Mesin (<i>Machine Learning</i>)	9
2.1.3. <i>Binary Classification</i>	11
2.1.4. <i>Gradient Boosting Classifier</i>	15
2.1.5. Diabetes	17
2.1.6. <i>AutoAI</i>	20
2.2. Penelitian Terdahulu.....	22
2.3. Kerangka Pemikiran	25
BAB III METODE PENELITIAN	27

3.1.	Desain Penelitian	27
3.2.	Metode Pengumpulan Data	29
3.3.	Analisis Data	30
3.4.	Metode Perancangan.....	32
3.4.1.	<i>Pre-Processing</i> Data.....	32
3.4.2.	Pembagian Dataset	33
3.4.3.	Pemilihan Model Otomatis.....	35
3.4.4.	Rekayasa Fitur Otomatis	41
3.4.5.	Optimasi <i>Hyperparameter</i>	42
3.5.	Lokasi dan Jadwal Penelitian	44
BAB IV	HASIL DAN PEMBAHASAN.....	45
4.1.	Hasil Penelitian.....	45
4.1.1.	Hasil Evaluasi Model	45
4.1.2.	Hasil Pengujian Data Tes	51
4.2.	Pembahasan	53
BAB V	SIMPULAN DAN SARAN	56
5.1	Simpulan.....	56
5.2	Saran	56
	DAFTAR PUSTAKA	58
	LAMPIRAN.....	61

DAFTAR GAMBAR

	Halaman
Gambar 2. 1 <i>Confusion Matrix</i>	13
Gambar 2. 2 IBM Cloud Pak For Data	21
Gambar 2. 3 Kerangka Pemikiran	25
Gambar 3. 1 Desain Peneilitian	27
Gambar 3. 2 Data yang tidak valid	33
Gambar 3. 3 <i>Data Training</i>	34
Gambar 3. 4 <i>Data Test</i>	34
Gambar 3. 5 <i>Input Data Training</i>	36
Gambar 3. 6 Menentukan kolom target.....	36
Gambar 3. 7 <i>Experiment Settings</i> (1).....	37
Gambar 3. 8 <i>Eksperiment Setting</i> (2).....	38
Gambar 3. 9 <i>Pipeline Leaderboard</i>	39
Gambar 3. 10 <i>Feature Summary</i>	42
Gambar 3. 11 <i>Enhancements</i>	43
Gambar 4. 1 Hasil <i>Confusion Matrix</i>	46
Gambar 4. 2 <i>ROC Curve</i>	48
Gambar 4. 3 <i>Precision Recall Curve</i>	49
Gambar 4. 4 <i>Model Evaluation Measure</i>	50
Gambar 4. 5 <i>Input Data Test</i>	51
Gambar 4. 6 Hasil <i>Output</i> Prediksi.....	52
Gambar 4. 7 Perbandingan Label <i>Outcome</i> dan Hasil Prediksi	53
Gambar 4. 8 Evaluasi Data Hasil Prediksi Positif.....	54

DAFTAR TABEL

	Halaman
Tabel 2. 1 Identifikasi Tipe Diabetes Berdasarkan Nilai Parameter	20
Tabel 3. 1 Analisis Berdasarkan Jenis Diabetes	31
Tabel 3. 2 Jadwal Penelitian	44

DAFTAR RUMUS

	Halaman
Rumus 2.1 <i>Accuracy</i>	14
Rumus 2.2 <i>Precision</i>	14
Rumus 2.3 <i>Recall</i>	14
Rumus 2.4 <i>F1-score</i>	14
Rumus 2.5 Fungsi Logistik	15
Rumus 2.6 Residual	15
Rumus 2.7 Koefisien Koreksi	16
Rumus 2.8 Pembaruan Prediksi	16
Rumus 2.9 Prediksi Akhir	17