

SeBats : Sistem Deteksi Penyakit Diabetes Berbasis Android

Dhavis Alvi Chandra¹, Nando Rahmat Prasetyo Mulyo², Donny Firdani³, Wahyu Cahyo Utomo⁴

^{1,2,3}Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Nusantara PGRI Kediri

E-mail: ¹dhavis.alv22@gmail.com, ²nando_nrpm@gmail.com, ³donnyfirdani@gmail.com,

⁴wahyu.utomo@unpkdr.ac.id

Abstrak – Diabetes merupakan kondisi degenerative kronis yang ditandai oleh produksi insulin yang tidak mencukupi di dalam pankreas. Ada berbagai faktor yang dapat memicu diabetes, termasuk faktor kehamilan, glukosa, tekanan darah, ketebalan kulit, insulin, BMI, silsilah diabetes atau keturunan, umur, dan faktor lainnya. Berdasarkan faktor tersebut, maka orang awam akan mengetahui gejala dasar penyakit diabetes. Diperlukannya diagnose dokter melalui pemeriksaan darah untuk memastikan apakah seseorang mengidap penyakit diabetes atau tidak. Penelitian ini bertujuan untuk membangun sebuah aplikasi deteksi Diabetes menggunakan algoritma Naïve Bayes berbasis Android. Hasil penelitian dapat bermanfaat untuk masyarakat umum dalam mendeteksi penyakit diabetes sejak dini. Tahapan penelitian yaitu mengumpulkan dataset diabetes, implementasi algoritma, dan evaluasi kinerja algoritma Naïve Bayes. Hasil penelitian menunjukkan bahwa algoritma Naïve Bayes dapat melakukan prediksi penyakit diabetes dengan kinerja yang baik. Nilai-nilai evaluasi kinerja algoritma Naïve Bayes untuk prediksi penyakit diabetes yaitu: *accuracy* 77%, *recall* 78%, *precision* 90%, dan *F1-score* 83%.

Kata Kunci — Naïve Bayes, Klasifikasi, Diabetes, Machine Learning, Android.

1. PENDAHULUAN

Diabetes Melitus (DM) adalah kondisi medis kronis yang ditandai oleh tingginya kadar glukosa dalam darah. Faktor genetic memainkan peran penting dalam DM, dengan hampir setengah dari semua penderita yang terkena penyakit ini memiliki faktor dari riwayat keluarga. Diabetes Melitus dibagi menjadi dua tipe utama, yaitu tipe 1 dan tipe 2. *Patogenesis* Diabetes Melitus tipe 1 (T1DM) terjadi ketika sel pankreas rusak dan gagal mengendalikan kadar glukosa darah secara efektif. Sedangkan, Diabetes Melitus tipe 2 (T2DM), juga dikenal sebagai DM yang tidak bergantung pada insulin, ditandai oleh resistensi insulin dan defisiensi sekresi insulin [1]. Karena gejala-gejala yang ditunjukkan oleh diabetes sering kali mirip dengan kondisi umum lainnya, banyak orang yang tidak menyadari bahwa mereka telah terkena penyakit ini dan bahkan mungkin sudah mengalami komplikasi. Oleh karena itu, diagnosis medis melalui pemeriksaan darah oleh dokter sangatlah penting untuk memastikan apakah seorang menderita diabetes atau tidak [2].

Terdapat beberapa penelitian telah dilakukan untuk mengembangkan model yang dapat memberikan hasil yang lebih akurat. Berdasarkan penelitian sebelumnya, telah ditemukan pendekatan yang lebih efektif dalam membangun model solusi untuk mendeteksi penyakit diabetes. Peneliti menggunakan tiga referensi penelitian sebelumnya yang dijadikan bahan rujukan untuk penelitian ini.

Penelitian yang pernah dilakukan dengan judul “IMPLEMENTASI DATA MINING KLASIFIKASI PENYAKIT DIABETES PADA PEREMPUAN MENGGUNAKAN NAÏVE BAYES” [3] yang dilakukan dengan pendekatan klasifikasi dengan kasus mengklasifikasikan penyakit diabetes pada perempuan agar membantu mendeteksi diabetes lebih awal, dengan hasil *accuracy* sebesar 78.50%, nilai *precision* sebesar 85.24%, nilai *recall* sebesar 83.64%, dan nilai AUC sebesar 0.8555, dengan mendapatkan hasil tersebut pada penelitian ini menyebutkan Naïve Bayes cocok digunakan untuk mengklasifikasikan penyakit diabetes.

Penelitian yang kedua dengan judul “Pendekatan *Data Science* untuk Deteksi Dini Diabetes Menggunakan Naïve Bayes Classifier” [4] dengan tujuan membuat sistem klasifikasi yang dapat melakukan pendeteksian dini terhadap penyakit diabetes. Penelitian tersebut menggunakan seratus data acak yang digunakan, dengan 80% dari data tersebut dialokasikan sebagai data *train* dan 20% sisanya sebagai data *test*. Berdasarkan evaluasi menggunakan *confusion matrix*, ditemukan bahwa pembagian yang paling ideal untuk *testing set* adalah 40%, dengan 60% sisanya digunakan sebagai *training set*. Dengan pembagian *dataset* ini, tingkat akurasi yang dicapai adalah sebesar 70%.

Penelitian yang ketiga dengan judul “Deteksi Dini Diabetes Mellitus pada Pasien Puskesmas Peneleh Menggunakan Naïve Bayes” [5] dengan tujuan membuat sistem deteksi penyakit diabetes. Penelitian tersebut

dirancang untuk mendukung tenaga medis dalam proses diagnose penyakit Diabetes Melitus pada pasien di Puskesmas Peneleh. Sistem ini memanfaatkan Algoritma *Naive Bayes* untuk deteksi dini penyakit dengan cara mengambil input berupa gejala dari pengguna dan menghitung probabilitas terjadinya penyakit Diabetes Melitus. tujuan utama dari sistem ini adalah untuk meningkatkan efisiensi proses diagnosis dan sistem ini telah menunjukkan tingkat keberhasilan sebesar 90% dalam pengujian sistem. Dari penelitian yang telah dilakukan dapat disimpulkan bahwa Algoritma *Naive Bayes* cocok digunakan untuk pendekatan klasifikasi penyakit diabetes dengan nilai akurasi yang cukup tinggi dan akan diaplikasikan ke dalam Android *Apps*.

2. METODE PENELITIAN

2.1 Metode Pengumpulan Data

Pengumpulan data dilakukan dengan menggunakan dataset dari kaggle yang bernama Pima Indians Diabetes yang berisikan data kesehatan atau tes laboratorium dengan label terkena diabetes ataupun tidak terkena diabetes.

Tabel 1. Contoh Dataset Rekam Medis Diabetes

<i>Pregnancies</i>	<i>Glucose</i>	<i>BloodPressure</i>	<i>SkinThickness</i>	<i>Insulin</i>	<i>BMI</i>	<i>DiabetesPedigree Function</i>	<i>Age</i>
6	148	72	35	0	33.6	0.627	50
1	85	66	29	0	26.6	0.351	31
8	183	64	0	0	23.3	0.672	32
1	89	66	23	94	28.1	0.167	21
0	137	40	35	168	43.1	2.288	33

Dari tabel 1 diketahui bahwa ada 8 variabel yang digunakan dalam penelitian, yaitu: *Pregnancies* (jumlah kehamilan), *Glucose* (kadar glukosa), *BloodPressure* (tekanan darah), *SkinThickness* (ketebalan kulit), *Insulin* (insulin), *BMI* (indeks massa tubuh), *DiabetesPedigreeFunction* (riwayat keturunan diabetes), *Age* (umur). Menurut penelitian yang dilakukan oleh Nofrian Deny Hendrawan, Mohammad Fadhol, Ahmad Herlyasa Sosro Pratama [6] bahwa pada tekanan darah 72 menunjukkan positif mengidap diabetes melitus, sedangkan tekanan darah 66 menunjukkan negatif atau tidak mengidap diabetes mellitus.

2.2 *Naive Bayes*

Klasifikasi merupakan suatu fungsi dalam data mining yang bertujuan untuk menghasilkan model yang dapat memprediksi kelas atau kategori dari objek dalam basis data. Proses ini terdiri dari dua tahap, yaitu tahap pembelajaran dan tahap pengklasifikasian. Pada tahap pembelajaran, algoritma klasifikasi akan membangun model klasifikasi dengan menganalisis data pelatihan. Tahap ini juga dapat dianggap sebagai tahap pembentukan fungsi atau pemetaan $Y=F(X)$ di mana Y adalah kelas hasil prediksi dan X adalah tuple yang kelasnya ingin diprediksi. Pada tahap pengklasifikasian, model yang telah dibangun akan digunakan untuk melakukan pengklasifikasian [7].

Metode yang digunakan dalam penelitian ini adalah *Naive Bayes*. *Naive Bayes* merupakan teknik klasifikasi probabilitas yang sederhana dengan menghitung sekelompok probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang telah diberikan. Algoritma ini menggunakan teorema *Bayes* dan mengasumsikan semua atribut independen atau tidak saling tergantung yang diberikan oleh nilai pada variabel kelas.

Bayes merupakan teknik prediksi probabilitas yang sederhana yang berdasarkan pada penerapan teorema *Bayes* (aturan *Bayes*) dengan asumsi independensi yang kuat (naif). Metode *Naive Bayes Classifier* menggunakan konsep probabilitas yang bertujuan untuk melakukan klasifikasi data pada kelas tertentu, metode *Naive Bayes Classifier* merupakan penyederhanaan dari *teorema Bayes*. Klasifikasi *Bayes* didasarkan pada *teorema Bayes* dengan formula sebagai berikut [8]:

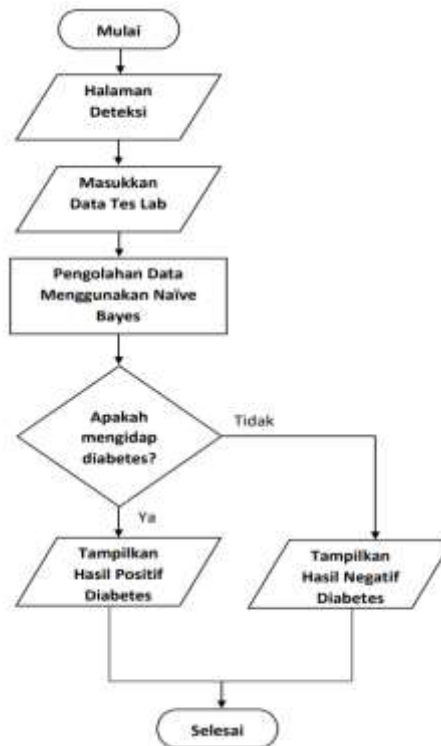
$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \dots\dots\dots(1)$$

Penjelasan dari formula (1) *teorema Bayes*, sebagai berikut:

- A = Hipotesis data B merupakan kelas spesifik.
- B = Data dengan kelas yang belum diketahui.
- $P(A|B)$ = Probabilitas akhir bersyarat (conditional probability) suatu hipotesis A terjadi jika diberikan bukti (evidence) B terjadi.
- $P(B|A)$ = Probabilitas sebuah bukti B terjadi akan pengaruh hipotesis A.
- $P(A)$ = Probabilitas awal (Priori) hipotesis A terjadi tanpa memandang bukti apapun.
- $P(B)$ = Probabilitas awal (Priori) bukti B terjadi tanpa memandang hipotesis/bukti lain.

2.3 Flowchart

Flowchart adalah diagram yang menggambarkan aliran kerja dalam suatu program atau prosedur sistem dengan struktur yang logis [9]. Flowchart dapat diaplikasikan untuk menggambarkan aktivitas manual, proses pemrosesan, atau keduanya. Flowchart terdiri dari serangkaian simbol yang digunakan untuk membangun struktur [10].



Gambar 1. Flowchart Sistem Deteksi

Pada gambar 1 merupakan gambaran alur dari aplikasi deteksi penyakit diabetes dengan menggunakan algoritma Naïve Bayes berbasis android. Diawali dengan muncul tampilan halaman form deteksi diabetes dan memasukkan semua data hasil tes lab. Jika sudah dimasukkan semua data, maka nanti akan muncul hasil orang tersebut mengidap diabetes atau tidak.

3. HASIL DAN PEMBAHASAN

3.1 Preprocessing Dataset

Data yang digunakan berasal dari website <https://kaggle.com>. Pada gambar 3 disajikan tangkapan layar *dataset Pima Indians Diabetes Database*.

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1
5	5	116	74	0	0	25.6	0.201	30	0
6	3	78	50	32	88	31.0	0.248	26	1
7	10	115	0	0	0	35.3	0.134	29	0
8	2	197	70	45	543	30.5	0.158	53	1
9	8	125	96	0	0	0.0	0.232	54	1

Gambar 2. Contoh Dataset Diabetes

Pada gambar 3 menunjukkan bahwa contoh 10 baris pertama dalam *dataset* tersebut. Pada penelitian ini menggunakan 8 variabel untuk mendeteksi penyakit diabetes. Untuk variabel yang digunakan adalah *Pregnancies*, *Glucose*, *Blood Pressure*, *Skin Thickness*, *Insulin*, *BMI*, *Diabetes Pedigree*, *Age* dan untuk variabel *outcome* digunakan sebagai variabel dependen. Untuk nilai 1 menunjukkan bahwa positif diabetes, sedangkan nilai 0 menunjukkan bahwa negatif diabetes.

Tabel 2. Konversi Data Numerik menjadi Kategorial

<i>Pregnancies</i>	<i>Glucose</i>	<i>BloodPressure</i>	<i>SkinThickness</i>	<i>Insulin</i>	<i>BMI</i>	<i>Diabetes Pedigree Function</i>	<i>Age</i>
Banyak Kehamilan	Tinggi	Normal	Normal	Rendah	Obesitas	Sedang	Lansia
Beberapa Kehamilan	Normal	Normal	Normal	Rendah	Berlebihan	Sedang	Dewasa
Banyak Kehamilan	Tinggi	Normal	Tipis	Rendah	Normal	Sedang	Dewasa
Beberapa Kehamilan	Normal	Normal	Normal	Normal	Berlebihan	Rendah	Dewasa
Beberapa Kehamilan	Tinggi	Normal	Normal	Tinggi	Obesitas	Tinggi	Dewasa
Banyak Kehamilan	Tinggi	Normal	Tipis	Rendah	Berlebihan	Sedang	Dewasa
Kehamilan Sedang	Normal	Normal	Normal	Normal	Obesitas	Sedang	Dewasa
Banyak Kehamilan	Tinggi	Normal	Tipis	Rendah	Obesitas	Rendah	Dewasa
Kehamilan Sedang	Tinggi	Normal	Tebal	Tinggi	Obesitas	Rendah	Lansia
Banyak Kehamilan	Tinggi	Normal	Tipis	Rendah	Kurus	Sedang	Lansia

Pada tabel 2 menunjukkan konversi dari data numerik menjadi kategorial. Berikut penjelasan tentang *range* kategori data pada tabel 2:

- Variabel *pregnancies* jika nilai lebih dari 3, maka akan dikategorikan memiliki banyak kehamilan.
- Variabel *glucose* jika nilai lebih dari 100, maka akan dikategorikan memiliki glukosa yang tinggi.
- Variabel *bloodpressure* jika nilai lebih dari 130, maka akan dikategorikan memiliki tekanan darah yang tinggi.
- Variabel *skintickness* jika nilai lebih dari 40, maka akan dikategorikan memiliki ketebalan kulit yang tebal.
- Variabel *insulin* jika nilai lebih dari 166, maka akan dikategorikan memiliki insulin yang tinggi.
- Variabel BMI jika nilai lebih dari 29.9, maka akan dikategorikan memiliki indeks masa tubuh yang atau berat badan yang obesitas.
- Variabel *diabetespedigreefunction* jika nilai lebih dari 0.7, maka akan dikategorikan memiliki riwayat keturunan diabetes yang tinggi.
- Variabel *age* jika nilai lebih dari 55, maka akan dikategorikan memiliki umur lansia.

3.2 Evaluasi Sistem

Setelah melakukan proses perhitungan dengan *Naïve Bayes* untuk melakukan klasifikasi pada setiap variabel, kemudian dilakukan perhitungan *confusion matrix* untuk mengetahui perhitungan akurasi model. Seperti pada gambar 3.

Kemudian dilakukan perhitungan untuk mengetahui hasil dari *accuracy*, *recall*, *precision*, *F-1 score*. Berdasarkan hasil dari *confusion matrix*.



Gambar 3. *Confusion Matrix*

Berdasarkan gambar 3 dapat dihitung *accuracy*, *recall*, *precesion*, *F-1 score*, sebagai berikut:

- a. $Accuracy = \frac{TP+TN}{TP+TN+FP+FN} = \frac{87+32}{87+32+10+25} = \frac{119}{154} = 0.77 = 77\%$
- b. $Recall = \frac{TP}{TP+FN} = \frac{87}{87+25} = 0.78 = 78\%$
- c. $Precision = \frac{TP}{TP+FP} = \frac{87}{87+10} = 0.90 = 90\%$
- d. $F-1\ score = \frac{TP}{TP+\frac{1}{2}(FP+FN)} = \frac{87}{87+\frac{1}{2}(10+25)} = 0.83 = 83\%$

Keterangan:

- True Positive* (TP) = Model memprediksi adanya serangan dan memang kenyataannya serangan terjadi.
- True Negative* (TN) = Model memprediksi tidak ada serangan dan memang kenyataannya tidak terjadi serangan.
- False Positive* (FP) = Model memprediksi adanya serangan, padahal tidak adanya serangan terjadi.
- False Negative* (FN) = Model memprediksi tidak adanya serangan, padahal serangan terjadi.

Dalam evaluasi model klasifikasi, terdapat *confusion matrix* penting yang dijelaskan dalam gambar 3. Pertama, *accuracy* merupakan ukuran seberapa sering model membuat prediksi yang benar, baik itu positif atau negatif. Dalam hal ini, akurasi dihitung dengan menjumlahkan *True Positive* (TP) dan *True Negative* (TN), lalu dibagi dengan total semua kasus. Kedua, *recall* merupakan ukuran seberapa baik model dalam mengidentifikasi kasus positif diabetes. *Recall* dihitung dengan membagi *True Positive* (TP) dengan jumlah *True Positive* (TP) dan *False Negative* (FN). Ketiga, *precision* merupakan ukuran seberapa baik model dalam melakukan prediksi positif diabetes. *Precision* dihitung dengan membagi (*True Positive*) TP dengan jumlah *True Positive* (TP) dan *False Positive* (FP). Terakhir, *F1-score* merupakan matrik yang mencoba menemukan keseimbangan antara *precision* dan *recall*. *F1-score* dihitung dengan menggabungkan *precision* dan *recall*.

3.3 Skenario Uji Coba Sistem

Pada tahap ini dilakukan skenario uji coba sistem dengan menghitung tingkat akurasi, dari sistem yang telah dibuat menggunakan persamaan *Confusion Matrix* [11]. Pada penelitian ini dilakukan tiga kali uji coba sistem. Uji coba pertama dilakukan menggunakan 20% *testing* dan 80% *data training* dari 768 data dengan 87 orang dinyatakan mengidap diabetes sesuai prediksi dan 25 orang dinyatakan mengidap diabetes namun hasilnya tidak mengidap diabetes, kemudian 32 orang tidak mengidap diabetes sesuai perkiraan, dan 10 orang diperkirakan tidak mengidap diabetes namun dinyatakan mengidap diabetes. Maka untuk model persamaan dapat dilihat pada tabel 3:

Tabel 3. Hasil Pengujian 20% *Data Testing* dan 80% *Data Training*

	<i>True Mengidap</i>	<i>True Tidak Mengidap</i>	<i>Class Precision</i>
Pred. Mengidap	87	25	90%
Pred. Tidak Mengidap	10	32	
<i>Class Recall</i>	78%		

Lalu uji coba kedua dengan menggunakan *data testing* 30% dan *data training* 70% dari 768 data dengan hasil 130 orang dinyatakan mengidap diabetes sesuai prediksi dan 34 orang dinyatakan mengidap diabetes namun hasilnya tidak mengidap diabetes, kemudian 47 orang tidak mengidap diabetes sesuai perkiraan, dan 22 orang diperkirakan tidak mengidap diabetes namun dinyatakan mengidap diabetes. Maka untuk model persamaan dapat dilihat pada tabel 4:

Tabel 4. Hasil Pengujian 30% *Data Testing* dan 70% *Data Training*

	<i>True Mengidap</i>	<i>True Tidak Mengidap</i>	<i>Class Precision</i>
Pred. Mengidap	130	34	85%
Pred. Tidak Mengidap	22	47	
<i>Class Recall</i>	79%		

Kemudian uji coba ketiga dengan menggunakan *data testing* 50% dan *data training* 50% dari 768 data dengan hasil 200 orang dinyatakan mengidap diabetes sesuai prediksi dan 75 orang dinyatakan mengidap diabetes namun hasilnya tidak mengidap diabetes, kemudian 75 orang tidak mengidap diabetes sesuai perkiraan, dan 49 orang diperkirakan tidak mengidap diabetes namun dinyatakan mengidap diabetes. Maka untuk model persamaan dapat dilihat pada tabel 5:

Tabel 5. Hasil Pengujian 50% *Data Testing* dan 50% *Data Training*

	<i>True Mengidap</i>	<i>True Tidak Mengidap</i>	<i>Class Precision</i>
Pred. Mengidap	200	58	80%
Pred. Tidak Mengidap	49	75	
<i>Class Recall</i>	78%		

Pengujian dilakukan sebanyak tiga kali dengan jumlah *data testing* dan *data training* yang berbeda yaitu, 20% *testing* dan 80% *training*, 30% *testing* dan 70% *training*, serta 50% *testing* dan 50% *training* dengan menggunakan *Confusion Matrix* yang mana terdiri dari *Precision*, *Recall*, *F1-Score*, serta *Accuracy* dilakukan pada dataset dengan menggunakan algoritma *Naïve Bayes* untuk menguji tingkat akurasi sistem dalam mendeteksi seseorang terkena diabetes atau tidak.

Dari hasil pengujian yang telah dilakukan, dengan tiga kali uji coba sistem, akurasi yang diperoleh dari data *test* 20% dan data *train* 80% adalah nilai *accuracy* sebesar 77%, *recall* sebesar 78%, *precision* sebesar 90%, *F1-score* sebesar 83%. Sementara pengujian menggunakan data *test* 30% dan *train* 70% menghasilkan *accuracy* sebesar 76%, *recall* sebesar 79%, *precision* sebesar 85%, *F1-score* sebesar 82% dan pengujian menggunakan data *test* 50% dan data *train* 50% menghasilkan *accuracy* sebesar 72%, *recall* sebesar 78%, *precision* sebesar 80%, *F1-score* sebesar 79%. Dari semua pengujian yang telah dilakukan, dapat diketahui bahwa pengujian menggunakan data *test* 20% dan data *train* 80% mempunyai nilai akurasi yang paling tinggi. Dengan akurasi yang lebih baik, maka sistem diaplikasikan ke dalam android *Apps*.

Dari nilai akurasi sistem dalam mengkalkulasi jumlah 3 dataset yang berbeda dan dengan masing - masing hasil akurasi yang berbeda. Hal tersebut diakarenakan sifat bawaan dari algoritma *Naïve Bayes* yang bersifat naif, yang mana tidak berlaku jika probabilitas kondisionalnya adalah 0 (nol), apabila nol maka probabilitas prediksi bernilai nol juga. Begitu pula sebaliknya, jika probabilitas bernilai 1 (satu), maka probabilitas prediksi bernilai satu juga [11].

3.4 Implementasi Algoritma *Naïve Bayes* dalam API

Pada langkah ini, dilakukan implementasi algoritma *Naïve Bayes* ke dalam API. API yang digunakan pada penelitian ini adalah Django. Berikut langkah-langkah implementasi algoritma ke dalam API Django:

- Import semua library yang dibutuhkan dalam proses implementasi algoritma.
- Melakukan *load* model algoritma *Naïve Bayes* sudah dilatih.
- Melakukan *request* dan menampilkan *response* API Django berupa JSON.
- Mengecek apakah hasil *response*, apakah model algoritma sudah berjalan atau belum.
- Jika API Django sudah berjalan dengan baik, maka akan menampilkan *response* sesuai model yang sudah dilatih, seperti gambar 4 berikut:

```

1  {
2  |   "data": {
3  |     "Pregnancies": 2.0,
4  |     "Glucose": 50.0,
5  |     "BloodPressure": 88.0,
6  |     "SkinThickness": 10.0,
7  |     "Insulin": 0.0,
8  |     "BMI": 34.0,
9  |     "DiabetesPedigreeFunction": 0.345,
10 |     "Age": 35.0
11 |   },
12 |   "hasil": "0",
13 |   "kesimpulan": "Pasien tidak terkena diabetes"
14 | }
    
```

Gambar 4. Contoh *Response* API Django

Pada gambar 4 menunjukkan bahwa setelah berhasil melakukan *request* API Django, maka *response* API Django akan menampilkan data sesuai request yang kita minta. Pada gambar 4 adalah contoh *response* dengan melakukan *request body* dalam bentuk JSON. Untuk tahap *request* API Django ini menggunakan aplikasi pihak ketiga yaitu Postman API.

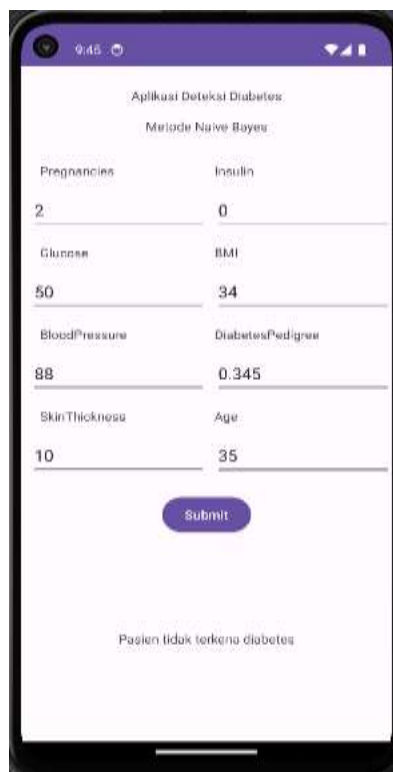
3.5 Tampilan UI Mobile Apps

Gambar 5 menampilkan interface dari bagian halaman deteksi yang terdapat delapan text field untuk *Pregnancies*, *Glucose*, *Blood Pressure*, *Skin Thickness*, *Insulin*, *BMI*, *Diabetes Pedigree*, *Age* dan sebuah button submit.



Gambar 5. Tampilan Halaman Deteksi

Gambar 6 merupakan hasil deteksi diabetes dari data yang dimasukkan pengguna aplikasi. Jika pengguna tersebut sudah mengisi semua *form* data tes lab yang dibutuhkan aplikasi, maka sistem aplikasi akan menampilkan apakah pengguna tersebut positif mengidap diabetes atau tidak.



Gambar 6. Tampilan Hasil Deteksi Diabetes

4. SIMPULAN

Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa sistem yang diuji dengan metode klasifikasi data mining *Naïve Bayes* menunjukkan variasi nilai akurasi berdasarkan proporsi data yang digunakan. Dalam tiga kali percobaan, ditemukan bahwa dengan proporsi *data testing* 20% dan *data training* 80%, akurasi yang dicapai adalah 77%. Sementara itu, dengan proporsi *data testing* 30% dan *data training* 70%, akurasi yang dicapai adalah 76%. Dan terakhir, dengan proporsi *data testing* 50% dan *data training* 50%, akurasi yang dicapai adalah 72%. Dengan demikian, dapat disimpulkan bahwa algoritma data mining dengan metode *Naïve Bayes* efektif dalam mendeteksi penyakit diabetes melitus berdasarkan kriteria faktor kehamilan, glukosa, tekanan darah, ketebalan kulit, insulin, BMI, silsilah diabetes atau keturunan, dan umur.

5. SARAN

Pada penelitian ini, peneliti menyadari masih ada kekurangan baik dari penulisan, sistem aplikasi, maupun teori yang digunakan dan mungkin bisa lebih dikembangkan lagi pada penelitian selanjutnya. Dalam penelitian mendatang, peneliti dapat mempertimbangkan untuk memanfaatkan metode klasifikasi data mining lainnya, seperti *K-Means*, *K-Nearest Neighbor*, *Support Vector Machine*, *Random Forest Classifier*. Dengan menerapkan berbagai metode, agar dapat mengetahui kelebihan dan kekurangan dari setiap metode dan menentukan metode mana yang menghasilkan nilai akurasi yang terbaik.

DAFTAR PUSTAKA

- [1] H. Wu, S. Yang, Z. Huang, J. He, and X. Wang, "Type 2 diabetes mellitus prediction model based on data mining," *Informatics Med. Unlocked*, vol. 10, no. August 2017, pp. 100–107, 2018, doi: 10.1016/j.imu.2017.12.006.
- [2] W. Apriliah, I. Kurniawan, M. Baydhowi, and T. Haryati, "SISTEMASI: Jurnal Sistem Informasi Prediksi Kemungkinan Diabetes pada Tahap Awal Menggunakan Algoritma Klasifikasi Random Forest," *J. Sist. Inf.*, vol. 10, pp. 163–171, 2021.
- [3] A. Veronica Agustin and A. Voutama, "Implementasi Data Mining Klasifikasi Penyakit Diabetes Pada Perempuan Menggunakan Naïve Bayes," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 2, pp. 1002–1007, 2023, doi: 10.36040/jati.v7i2.6808.
- [4] N. Ningsih, A. Aprianto, and A. Angeline, "Pendekatan Data Science untuk Deteksi Dini Diabetes Menggunakan Naive Bayes Classifier," *J. Inf. Syst. Hosp. Technol.*, vol. 5, no. 1, pp. 26–31, 2023, doi: 10.37823/insight.v5i1.300.
- [5] N. Fuad, "Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika," pp. 219–224, 2021.
- [6] N. D. Hendrawan, M. Fadhol, A. Herlyasa, and S. Pratama, "Pengembangan Model Machine Learning menggunakan Docker dan Flask Pada Python sebagai Deteksi Diabetes Melitus dari Pedigree Function dan Tekanan Darah," no. November, pp. 372–378, 2022.
- [7] Heliyanti Susana, "Penerapan Model Klasifikasi Metode Naive Bayes Terhadap Penggunaan Akses Internet," *J. Ris. Sist. Inf. dan Teknol. Inf.*, vol. 4, no. 1, pp. 1–8, 2022, doi: 10.52005/jursistekni.v4i1.96.
- [8] R. Antoni, "MELLITUS DENGAN MENGGUNAKAN PENDEKATAN NAÏVE BAYESIAN BERBASIS WEB (Studi Kasus : PUSKESMAS Kelurahan Grogol 3)," vol. 2, no. 1, pp. 334–348, 2023.
- [9] R. A. Y. Manurung and A. D. Manuputty, "Perancangan Sistem Informasi Lembaga Kemahasiswaan Universitas Kristen Satya Wacana Salatiga," *J. SITECH Sist. Inf. dan Teknol.*, vol. 3, no. 1, pp. 9–20, 2020, doi: 10.24176/sitech.v3i1.4703.
- [10] Q. Budiman, S. Mouton, L. Veenhoff, and A. Boersma, "ANALISIS PENGENDALIAN MUTU DI BIDANG INDUSTRI MAKANAN (Studi Kasus: UMKM Mochi Kaswari Lampion Kota Sukabumi)," *J. Inov. Penelit.*, vol. 1, no. 0.1101/2021.02.25.432866, pp. 1–15, 2021.
- [11] R. A. Iswanto, J. Sahertian, and M. A. D. Widyadara, "Pengembangan Sistem Prediksi Kelulusan Mahasiswa Berdasarkan Nilai Disiplin, Hasil Belajar, Aktivitas Sosial Ekonomi, dan Aktivitas Organisasi Menggunakan Algoritma Naïve Bayes," *Pros. Semin. Nas. Teknol. Dan Sains*, vol. 1, pp. 349–358, 2022.