



# IMPLEMENTASI METODE TF-IDF DAN ALGORITMA NAIVE BAYES DALAM APLIKASI DIABETIC BERBASIS ANDROID

**I Wayan Alston Argodi**

Program Studi Informatika, Fakultas Ilmu Komputer  
Universitas Pembangunan Nasional “Veteran” Jawa Timur

**Eva Yulia Puspaningrum**

Program Studi Informatika, Fakultas Ilmu Komputer  
Universitas Pembangunan Nasional “Veteran” Jawa Timur

**Muhammad Muharrom Al Haromainy**

Program Studi Informatika, Fakultas Ilmu Komputer  
Universitas Pembangunan Nasional “Veteran” Jawa Timur

Korespondensi penulis: [19081010163@student.upnjatim.ac.id](mailto:19081010163@student.upnjatim.ac.id)

**Abstract.** *Diabetes is a serious disease that occurs when the pancreas does not produce enough insulin as a hormone that regulates blood sugar in the body. This disease also has an impact on health. This research builds an Android-based application called Diabetic to help classify and provide information related to diabetes and analyze the performance of the Term Frequency Inverse Document Frequency method and the Naive Bayes algorithm. The Term Frequency Inverse Document Frequency method is a technique for calculating the presence of words in a collection of documents by creating document vectors. The Naive Bayes algorithm is an algorithm that uses probability to solve a classification case. This algorithm has an efficient and fast calculation. Based on this research, it is known that the Naive Bayes Algorithm produces an accuracy of 66% by taking a computation time of 39 seconds with a memory consumption of 80 to 351 mb.*

**Keywords:** *Diabetes, Term Frequency Inverse Document Frequency, Naive Bayes algorithm, Android apps*

**Abstrak.** Diabetes merupakan penyakit serius yang terjadi ketika pankreas tidak menghasilkan cukup Insulin sebagai hormon yang mengatur gula darah dalam tubuh penyakit ini juga berdampak pada kesehatan, penelitian ini membangun aplikasi berbasis Android bernama Diabetic membantu dalam klasifikasi dan informasi terkait diabetes dan menganalisa performa metode Term Frequency Inverse Document Frequency serta algoritma Naive Bayes. Metode Term Frequency Inverse Document Frequency merupakan teknik menghitung kehadiran kata pada kumpulan dokumen dengan membuat vektor dokumen Algoritma Naive Bayes merupakan algoritma yang menggunakan probabilitas dalam menyelesaikan sebuah kasus klasifikasi algoritma ini memiliki perhitungan yang efisien dan cepat. Berdasarkan penelitian ini diketahui Algoritma Naive Bayes menghasilkan akurasi sebesar 66% dengan menempuh waktu komputasi selama 39 detik dengan konsumsi memori sebesar 80 sampai dengan 351 mb.

**Kata kunci:** Diabetes, *Term Frequency Inverse Document Frequency*, Algoritma Naive Bayes, Aplikasi Android

## LATAR BELAKANG

Diabetes merupakan penyakit serius yang terjadi ketika pankreas tidak menghasilkan cukup Insulin sebagai hormon yang mengatur gula darah. Jumlah kasus diabetes terus meningkat selama beberapa dekade terakhir (infoDatin,2020). menurut (Fountaine,dkk,2019) teknologi dapat berperan penting dalam bidang kesehatan untuk membangun alat pendukung keputusan yang didukung teknologi dapat membantu dalam mendiagnosa dan mengobati penyakit dengan lebih baik seperti aplikasi seluler dapat memberikan manajemen penyakit. Beberapa peneliti didunia mengusulkan sistem prediksi penyakit kesehatan menggunakan Machine Learning oleh (Alam,,dkk.2019) melakukan penelitian untuk klasifikasi penyakit diabetes menggunakan algoritma Random Hasil pengujian algoritma Random Forest menghasilkan akurasi berbeda pada setiap pemberian nilai parameter yang berbeda dengan jumlah fitur lebih banyak 30 fitur akurasi yang dihasilkan sebesar 75.30% sedangkan nilai fitur lebih rendah 10 dengan nilai sampel sebesar 340 menghasilkan akurasi sebesar 98%, selanjutnya penelitian oleh (Sneha.dkk,2019) Hasil pengujian menunjukan algoritma Random Forest menghasilkan akurasi sebesar 75.39% dengan ketepatan data tidak diabetes 89 dan benar diabetes 10 hasil class *precision* 89.90% dan 73.24%.dan algoritma klasifikasi Naive Bayes menghasilkan akurasi sebesar 73.48% dengan ketepatan data tidak diabetes 49 dan benar diabetes 30 hasil class *precisison* 62.03% dan 79.47% Penelitian ini membangun aplikasi bernama diabetic dan menganalisis untuk mengetahui bagaimana performa dan akurasi metode *Term Frequency Inverse Document Frequency* serta algoritma Naive Bayes dalam aplikasi diabetic berbasis Android menerapkan komputasi secara *on device machine learning*. Penelitian ini mengembangkan dan menganalisa aplikasi diabetic berbasis Android menggunakan metode *Term Frequency Inverse Document Frequency* dan algoritma Naive Bayes dalam aplikasi berbasis android sistem prediksi secara langsung atau realtime sehingga proses perhitungan hanya dalam lingkup aplikasi *Native Android*.

## KAJIAN TEORITIS

### A. Penelitian sebelumnya

Pada penelitian oleh (Sneha.dkk,2019) berjudul melakukan penelitian terkait deteksi awal deteksi penyakit diabetes menggunakan seleksi fitur optimal, penelitian ini menggunakan algoritma Naive Bayes, Random Forest, dan Support Vector Machine dengan dataset diabetes dataset terdapat 2500 item data dengan 15 atribut. Dataset selanjutnya melalui proses *pre-processing* dengan menyeleksi atribut yang terbaik proses selanjutnya dataset dimasukkan kedalam algoritma untuk dilakukan proses pelatihan dengan dataset. Hasil pengujian menunjukan algoritma Support Vector Machine menghasilkan akurasi sebesar 77.37%,kemudian algoritma Random Forest menghasilkan akurasi sebesar 75.39% dan algoritma klasifikasi Naive Bayes menghasilkan akurasi sebesar 73.48%. Selanjutnya oleh (Alam.dkk,2019) yang berjudul “*A model for early prediction of diabetes*” melakukan penelitian implementasi model based untuk prediksi diabetes dengan menggunakan algoritma *Artificial Neural Network* dan Random Forest dengan 500 rekam dataset diabetes melalui proses *preprocessing* untuk *cleaning dataset* mencari data dan nilai hilang dalam dataset, *data reduction*, dan *data transformation*.Pengujian pada penelitian ini dilakukan dengan berbagai jumlah pohon keputusan dalam algoritma Random Forest menghasilkan akurasi berbeda pada setiap nilai jumlah, hasil pengujian menunjukan jumlah pohon keputusan 8 menghasilkan akurasi 74.3% dan kemudian dinaikan menjadi 50 pohon keputusan menghasilkan akurasi sebesar 74.6%. (Fitriyani.dkk,2019) melakukan penelitian pengembangan *Model Based* untuk prediksi penyakit diabetes dan hipertensi. Terdapat 4 jenis dataset yang digunakan dalam penelitian ini dataset pertama memiliki atribut lebih seperti tekanan darah kemudian *dataset* dua dan tiga memiliki kesamaan dalam atribut, selanjutnya *dataset* empat memiliki atribut umur, hipertensi serta tekanan darah., hasil pengujian menunjukan algoritma Random Forest mendapatkan akurasi sebesar 69.9%, 56.96%, 49.96% dan 72.48%.Penelitian sebelumnya sistem prediksi diterapkan pada Model- Based sehingga belum diterapkan secara

implementasi untuk sistem klasifikasi, dengan hal ini algoritma dapat berjalan dengan baik pada dengan komputasi tinggi.

## B. Algoritma Naive Bayes

Algoritma Naive Bayes merupakan algoritma menggunakan nilai probabilitas akhir untuk menyelesaikan suatu kasus klasifikasi (Irwansyah,2021) berdasarkan teorema Bayes untuk menentukan asumsi prediksi, algoritma ini menggunakan dataset sebagai input kemudian analisa performa dan prediksi kelas berdasarkan label menggunakan Bayes, Naive Bayes dapat digunakan untuk dataset dengan jumlah yang cukup besar (Tripathi,2019) Pengklasifikasi Bayesian menetapkan kelas yang paling mungkin untuk menjadi vektor fitur dengan fitur kelas independen yang diberikan, Naive Bayes memiliki optimalitas dalam nol dan satu atau kesalahan klasifikasi belum tentu terkait dengan kualitas kecocokan dengan distribusi probabilitas pengklasifikasi optimal diperoleh selama distribusi aktual dan perkiraan setuju pada kelas yang paling mungkin (Tripathi,2019).

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)} \quad (1)$$

Pada persamaan 1 merupakan perhitungan algoritma Naive Bayes, X merupakan data class yang belum diketahui kemudian c adalah hipotesis data suatu class spesifik kemudian  $P(C|X)$  merupakan probabilitas hipotesis berdasarkan kondisi atau *posteriori probability* selanjutnya  $P(c)$  merupakan probabilitas hipotesis,  $P(X|C)$  merupakan probabilitas berdasarkan kondisi hipotesis kemudian  $P(x)$  merupakan probabilitas c.

## C. Confusion Matrix

*Confusion matrix* merupakan cara untuk melakukan sebuah evaluasi terhadap klasifikasi. Matrix ini berisikan data target prediksi yang dibandingkan dengan data target aktual (Irwansyah,2021), Accuracy atau rasio prediksi benar memberikan gambaran terkait seberapa akurat sebuah model Machine Learning memprediksi nilai dengan benar dengan keseluruhan data.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (2)$$

Pada persamaan 2 merupakan rumus akurasi dalam mengukur akurasi dan seberapa benar prediksi model, dalam gambar diatas memiliki aturan *true positive* (TP), *true negatif* (TN), *false positive* (FP) dan *false negatif* (FN). *Precision* merupakan rasio prediksi benar positif dengan keseluruhan hasil prediksi yang positif berdasarkan jumlah kelas yang benar positif (Irwansyah,2021) ,berikut ini persamaan untuk menghitung nilai Precision :

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3)$$

Pada persamaan 3 merupakan rumus Precision dengan menghitung jumlah *true positif* dibagi dengan jumlah antara *true positive* dibagi dengan nilai *false positive*, Recall memberikan bagaimana keberhasilan model dalam menemukan kembali sebuah informasi, Recall memiliki rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

Pada persamaan 4 merupakan rumus untuk menghitung Recall dengan nilai *true positive* dibagi dengan jumlah *true positive* dan *false*, Specify merupakan perhitungan untuk mencari kebenaran memprediksi negatif dibandingkan dengan keseluruhan data negatif dalam data dan prediktor.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (5)$$

Pada persamaan 5 merupakan rumus precision dengan menghitung jumlah *true negative* dibagi dengan jumlah antara *true negative* dibagi dengan nilai *false positive*.F1 Score merupakan perbandingan rata-rata *precision* dan *recall* yang telah dibobotkan.

$$F1 \text{ Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

Pada persamaan 6 merupakan rumus untuk mencari F1 Score dengan pembobotan 2 kali nilai Recall dikalikan dengan Precision kemudian dibagi dengan jumlah Recall dan Precision (Irwansyah,2021).

## METODE PENELITIAN

### A. Tahapan Penelitian

Tahapan penelitian merupakan alur dari penelitian ini berjalan dari awal hingga akhir.



Gambar 1. Alur penelitian

Pada gambar 1 diatas merupakan diagram alur dari penelitian ini dimulai dengan studi literatur terkait referensi dan wawasan yang diperlukan untuk mendukung penelitian ini selanjutnya akusisi dataset yang digunakan dalam penelitian dataset gambar selanjutnya akan melalui pre-processing untuk mengurangi outlier. Tahap selanjutnya yaitu analisa kebutuhan meliputi perangkat keras dan perangkat lunak untuk menunjang penelitian. Tahap selanjutnya yaitu perancangan dan implementasi perancangan meliputi pembuatan diagram rancangan aplikasi seperti perancangan fungsi *Term Frequency Inverse Document Frequency* dan algoritma Naive Bayes kemudian implementasi pada pembuatan aplikasi Android. Tahap selanjutnya yaitu pengujian meliputi pengujian aplikasi Android.

### B. Dataset Diabetes

Pada tahap ini yaitu akusisi terkait dataset diabetes yang digunakan untuk klasifikasi dalam aplikasi ini, dataset ini diakusisi dari situs open source data kaggle dengan pemilik data UCI Machine Learning terdapat sembilan atribut dalam dataset.

Tabel 1. Dataset Diabetes

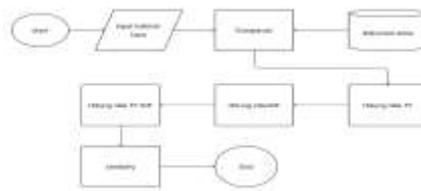
Pregnancies	Glucose	Blood Pressure	Skin Thickness	Insulin	BMI	pedigree	umur
-------------	---------	----------------	----------------	---------	-----	----------	------

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
5	116	74	238	3	2	50.2	2

Pada tabel 1 diatas merupakan Pima Diabetes Database terdapat beberapa fitur sebagai indikator riwayat kesehatan diabetes seperti kehamilan, jumlah glukosa dalam darah, tekanan darah, ketebalan kulit, pemberian insulin, berat badan, diabetes pedigree, umur, dan indikator diabetes terkena ya sebagai 1 dan tidak sebagai 0. Dataset ini berjumlah 768 data yang kemudian akan dibagi menjadi sebagai data latih dan data uji.

### C. Term Frequency Inverse Document Frequency

Ekstraksi dan pemilihan fitur penting untuk Chatbot pada penelitian ini, tahap ini merupakan pembuatan fungsi *Term Frequency Inverse Document Frequency* untuk mencari kesamaan antara dokumen baru dengan dokumen data lama pada database.



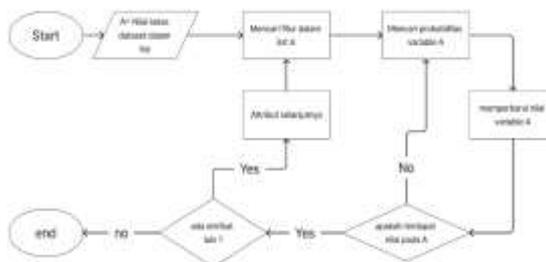
**Gambar 2.** alur Term Frequency Inverse Document Frequency

Pada gambar 2 diatas merupakan alur dari perhitungan *Term Frequency Inverse Document Frequency* kalimat baru yang telah melalui proses tokenisasi selanjutnya akan diambil dengan dokumen kalimat lama dalam database.kalimat yang telah menjadi token akan dibanding kemunculan untuk mencari nilai *Term Frequency* kemudian untuk mencari nilai *Inverse Document Frequency* jumlah dokumen yang dibandingkan dibagi dengan dokumen yang memiliki ketentuan.

### D. Algoritma Naive Bayes

Algoritma Naive Bayes merupakan algoritma menggunakan nilai probabilitas untuk menyelesaikan suatu kasus klasifikasi dengan berdasarkan teorema Bayes untuk menentukan asumsi prediksi, algoritma ini menggunakan dataset sebagai input kemudian analisa performa dan prediksi kelas berdasarkan label

menggunakan Bayes, secara teori algoritma ini menghitung probabilitas kelas input dan melakukan prediksi terhadap kelas baru atau yang belum diketahui.



Gambar 3. Algoritma Naive Bayes

Pada gambar 3 diatas merupakan diagram alur dari algoritma Naive Bayes. Algoritma Naive Bayes menggunakan nilai probabilitas atau kemungkinan dalam menyelesaikan sebuah kasus, dengan menggunakan data baru dan data lama dalam mencari probabilitas data baru seperti pada pembahasan sebelumnya Naive Bayes merupakan algoritma menggunakan nilai probabilitas untuk menyelesaikan suatu kasus klasifikasi dengan berdasarkan teorema Bayes untuk menentukan asumsi prediksi, algoritma ini menggunakan dataset sebagai input kemudian analisa performa dan prediksi kelas berdasarkan label menggunakan Bayes, secara teori algoritma ini menghitung probabilitas kelas input dan melakukan prediksi terhadap kelas baru atau yang belum diketahui.

### HASIL DAN PEMBAHASAN

Pengujian atau evaluasi ini dilakukan untuk mengetahui akurasi algoritma yang telah dilatih sebelumnya berdasarkan performa dan akurasi dari model pada aplikasi menggunakan *confusion matrix* dan performa waktu komputasi metode Term Freunecy Inverse Document Frequency dan algoritma Naive Bayes.

Tabel 2. Data Pengujian

Skenario	Persentase data latih	Jumlah Data Latih	Persentase Data uji	Jumlah data uji
1	90%	691	10%	71
2	80%	613	20%	154
3	70%	536	30%	231
4	60%	460	40%	308
5	50%	383	50%	383
6	40%	460	60%	308

Pada tabel 2 merupakan data untuk pengujian klasifikasi diabetes untuk digunakan dalam algoritma Naive Bayes terdapat 691 data pelatihan dan 77 data untuk pengujian algoritma dengan menghasilkan dua class yaitu ya dan tidak termasuk diabetes. jumlah data

pengujian digunakan dalam penelitian ini pada algoritma Naive Bayes dengan ini dapat diketahui perbandingan kedua algoritma tersebut dalam aplikasi diabetic.

**Tabel 3. Pengujian algoritma Naive Bayes**

Skenario	akurasi	presisi	recall	specify
1	61%	36%	33%	83%
2	61%	39%	27%	85%
3	62%	38%	21%	89%
4	65%	48%	10%	95%
5	66%	48%	25%	87%
6	64%	51%	26%	85%

Pada tabel 3 merupakan hasil dari pengujian algoritma Naive Bayes dengan beberapa skenario setiap skenario menggunakan jumlah data latih dan data uji yang berbeda skenario pertama menggunakan 90% data latih dan 10% data uji, hasil pengujian menunjukan skenario lima dengan 50% data latih dan 50% data uji menghasilkan akurasi tertinggi, pada pengujian ini jika algoritma Naive Bayes memiliki data latih semakin sedikit dibandingkan data uji maka akurasi bertambah, namun pada skenario enam akurasi berkurang menjadi 64%.

**Tabel 4. Penggunaan sumber daya algoritma Naive Bayes**

Skenario	Waktu komputasi	Pengunaan RAM	Penggunaan CPU
1	12 detik	80-112 MB	29-70%
2	15 detik	88-157 MB	28-73%
3	18 detik	80-185 MB	26-72%
4	20 detik	90-230 MB	25-68%
5	39 detik	80-351 MB	24-72%
6	48 detik	90-335 MB	25-70%

Pada tabel 4 merupakan jumlah sumber daya yang digunakan dalam proses klasifikasi algoritma Naive Bayes dengan menggunakan 383 data latih serta 231 data uji, berdasarkan tabel hasil confusion matrix sebelumnya, algoritma menghasilkan akurasi sebesar 66% dengan menempuh waktu komputasi sebesar 39 detik serta penggunaan memori sebanyak 80 sampai dengan 351 mb dan penggunaan cpu sebesar 24 sampai dengan 72%.

## **KESIMPULAN DAN SARAN**

### **A. Kesimpulan**

Berdasarkan implementasi dan pengujian, aplikasi diabetic berhasil menjawab pertanyaan terkait diabetes dengan *Term Frequency Inverse Document Frequency* dan melakukan klasifikasi diabetes dengan menggunakan algoritma Naive Bayes. Algoritma Naive Bayes memiliki keunggulan dalam waktu komputasi. *Term Frequency Inverse Document Frequency* menempuh waktu selama 1.39 detik akurasi metode ini dalam pengujian pertanyaan mendapatkan beberapa jawaban yang tidak sesuai karena dalam proses perhitungan melakukan pembobotan pada kalimat dalam database sehingga jika pertanyaan dengan kalimat yang memiliki struktur sama maka dianggap kalimat tersebut sama dan memiliki nilai lebih besar jika kalimat tersebut berulang kali muncul. algoritma Naive Bayes mendapatkan akurasi 66% dengan performa waktu komputasi selama 39 detik.

### B. Saran

Saran yang dapat diberikan untuk penelitian selanjutnya penambahan variasi dataset diabetes, algoritma Naive Bayes dalam prosesnya membutuhkan lebih banyak pengetahuan awal karena tingkat keberhasilan bergantung pada pengetahuan awal tersebut. Penambahan optimasi Laplace Smoothing algoritma Naive Bayes untuk meningkatkan akurasi dan mencegah jumlah pertumbuhan probabilitas nol berdasarkan pengujian sebelumnya algoritma ini menghasilkan akurasi sebesar 60%.

### DAFTAR REFERENSI

- KementrianKesehatanInfodatin(2020).InfodatinDiabetesmiletus<https://www.kemkes.go.id/article/view/20120100005/infodatin-tetap-produktif-cegah-dan-atasi-diabetes-melitus-2020.html>
- Alam, Talha Mahboob, et al. "A model for early prediction of diabetes." *Informatics in Medicine Unlocked* 16 (2019): 100204.
- Fountaine, T., Lembong, J., & Nair, R. (2016). Tackling Indonesia's diabetes challenge: Eight approaches from around the world. McKinsey Global Institute.
- International Diabetes Forum (2020) Diabetes Atlas 10th edition.
- Sneha, N., & Gangil, T. (2019). Analysis of diabetes mellitus for early prediction using optimal features selection. *Journal of Big data*, 6(1), 1-19.
- Tripathi, A., Yadav, S., & Rajan, R. (2019, July). Naive Bayes classification model for the student performance prediction. In 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT) (Vol. 1, pp. 1548-1553). IEEE.
- Kjos, Andrea L., Amy G. Vaughan, and Anuj Bhargava. "Impact of a mobile app on medication adherence and adherence-related beliefs in patients with type 2 diabetes." *Journal of the American Pharmacists Association* 59.2 (2019): S44-S51.
- Rish, I. (2001, August). An empirical study of the naive Bayes classifier. In IJCAI 2001 workshop on empirical methods in artificial intelligence (Vol. 3, No. 22, pp. 41-46).
- Thongyoo, P., Anantapanya, P., Jamsri, P., & Chotipant, S. (2020). A Personalized Food

Recommendation Chatbot System for Diabetes Patients. In Cooperative Design, Visualization, and Engineering: 17th International Conference, CDVE 2020, Bangkok, Thailand, October 25–28, 2020, Proceedings 17 (pp. 19-28). Springer International Publishing.

Fitriyani, N. L., Syafrudin, M., Alfian, G., & Rhee, J. (2019). Development of disease prediction model based on ensemble learning approach for diabetes and hypertension. *Ieee Access*, 7, 144777-144789.

Al Haromainy, M. M., Setyawan, D. A., Waluya, O. K., & Arifin, A. Z. (2019). Ambiguitas Machine Translation pada Cross Language Chatbot Bea Cukai. *Register: Jurnal Ilmiah Teknologi Sistem Informasi*, 5(1), 55-62.

Onah, J. O., Abdullahi, M., Hassan, I. H., & Al-Ghusham, A. (2021). Genetic Algorithm based feature selection and Naïve Bayes for anomaly detection in fog computing environment. *Machine Learning with Applications*, 6, 100156.

Abbasi, M., & Martins, P. (2020). An Overview of Search and Match Algorithms Complexity and Performance. In *Bioinformatics and Biomedical Engineering: 8th International Work-Conference, IWBBIO 2020, Granada, Spain, May 6–8, 2020, Proceedings 8* (pp. 457-471). Springer International Publishing.

Islam, M. T., Raihan, M., Farzana, F., Aktar, N., Ghosh, P., & Kabiraj, S. (2020, July). Typical and non-typical diabetes disease prediction using random forest algorithm. In *2020 11th International conference on computing, communication and networking technologies (ICCCNT)* (pp. 1-6). IEEE.

Kadhim, A. I. (2019, April). Term weighting for feature extraction on Twitter: A comparison between BM25 and TF-IDF. In *2019 international conference on advanced science and engineering (ICOASE)* (pp. 124-128). IEEE.

Phetkrachang, K., & Kittiphattanabawon, N. (2020). Fuzzy TF-IDF Weighting in Synonym for Diabetes Question and Answers. In *Recent Advances in Information and Communication Technology 2019: Proceedings of the 15th International Conference on Computing and Information Technology (IC2IT 2019)* 15 (pp. 59-68). Springer International Publishing.

Kotlin: Statically typed programming language for the JVM, Android and the browser. JetBrains.(n.d.).RetrievedApril17,2023,from<https://www.jetbrains.com/opensource/kotli/>

Irwanysah., 2021. Machine Learning untuk pemula. " Penerbit Informatika."

Géron, A. (2022). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow. " O'Reilly Media, Inc."