

Pengembangan Model *Machine Learning* menggunakan Docker dan Flask Pada Python sebagai Deteksi Diabetes Melitus dari *Pedigree Function* dan Tekanan Darah

Nofrian Deny Hendrawan¹, Mohammad Fadhol², Ahmad Herlyasa Sosro Pratama³

^{1,2}Fakultas Teknologi Informasi, Universitas Merdeka Malang, ³Fakultas Ekonomi dan Bisnis, Universitas Merdeka Malang

nofrian.hendrawan@unmer.ac.id¹, mohammad.fadhol@unmer.ac.id², ahmad.pratama@unmer.ac.id³

Abstract

In recent years, Machine Learning and Deep Learning have been used to predict diabetes and its complications. Of the several factors associated with a person being diagnosed with diabetes we focused on predicting the features of diabetes classification based on pedigree function and blood pressure. To get better validation results, we try to combine the two as a diabetes detection feature. And to make it even easier to implement detection systems, docker is a containerization service that enables convenient deployment of websites, databases, application APIs, and machine learning models (ML). From the results of using Docker, we found a system that is more efficient in terms of running time apps. And the results of our classification use 2 Machine Learning classification methods, namely K-NN and Naïve Bayes, each of which produces > 60% accuracy from training and testing data. This shows that the 2 features we chose as classification attributes can be used as a reference for diabetes detection.

Keywords: *Diabetes Detection, Pedigree Function, Blood Pressure, Machine Learning, K-NN, Naïve Bayes*

Abstrak

Beberapa tahun terakhir, *Machine Learning* dan *Deep Learning* telah digunakan untuk memprediksi diabetes dan komplikasinya. Dari beberapa faktor yang terkait dengan seseorang dinyatakan diabetes kami memfokuskan untuk memprediksi fitur klasifikasi diabetes berdasarkan *pedigree function* dan tekanan darah. Untuk lebih mendapatkan hasil validasi yang lebih baik, maka kita mencoba menggabungkan keduanya sebagai fitur deteksi diabetes. Dan untuk lebih memudahkan dalam implementasi sistem deteksi, docker adalah layanan containerisasi yang memungkinkan kenyamanan penerapan situs web, database, API aplikasi, dan mesin model pembelajaran (ML). Dari hasil penggunaan docker ini kami mendapati sistem yang lebih efisien dari segi *running time apps*. Metode klasifikasi, yaitu menggunakan *Machine Learning*. Dengan menggunakan 3 fitur data yang kami ubah dari Kaggle, 3 fitur data ini terdiri dari Blood Pressure, Diabetes Pedigree Function, dan Outcome untuk hasil deteksinya. Ini menunjukkan jika 3 fitur yang kami pilih sebagai atribut klasifikasi dapat dijadikan rujukan deteksi diabetes.

Kata kunci: *Deteksi Diabetes, Pedigree Function, Tekanan Darah, Machine Learning*

1. Pendahuluan

Diabetes melitus merupakan salah satu penyakit kronis yang belum dapat disembuhkan. Menurut World Health Organization (WHO), prevalensi penderita diabetes di Indonesia beradapada peringkat ke tujuh di dunia dengan estimasi penderita ± 10 juta orang. Diabetes melitus disebabkan oleh produksi insulin sedikit dari pankreas atau tubuh tidak dapat memanfaatkan insulin yang dihasilkan secara efektif (Kemenkes, 2018).

Gejala paling umum pada penderita diabetes melitus yaitu tingginya kadar gula dalam darah. Beberapa penderita diabetes di Indonesia mengalami komplikasi. Peluang mengalami komplikasi disebabkan akibat beberapa penderita diabetes yang tidak mengetahui bahwa dirinya menderita penyakit tersebut serta terlambat dalam penanganan penyakitnya. tidak mengetahui bahwa dirinya menderita penyakit tersebut serta berpeluang mengalami komplikasi jika terlambat penanganannya.

Tes gula darah merupakan pemeriksaan diabetes melitus yang pada umumnya dilakukan saat ini. Pemeriksaan tersebut masih dilakukan secara invasif atau masih dibutuhkan sampel darah dengan cara menusuk jarum ke jari pasien. Jangka waktu yang lama merupakan salah satu kendala dalam mendapatkan hasil pemeriksaan tersebut. Selain itu, biaya yang dibutuhkan untuk pemeriksaan juga cukup tinggi.

Berdasarkan permasalahan tersebut, penelitian ini membantu memprediksi diabetes dan komplikasinya dengan menggunakan *Machine Learning* dan *Deep Learning*. Dari beberapa faktor yang terkait dengan seseorang dinyatakan diabetes kami memfokuskan untuk memprediksi fitur klasifikasi diabetes berdasarkan *pedigree function* dan tekanan darah.

Untuk lebih mendapatkan hasil validasi yang lebih baik, maka kita mencoba menggabungkannya sebagai fitur deteksi diabetes. Dan untuk lebih memudahkan dalam implementasi sistem deteksi, docker adalah layanan containerisasi yang memungkinkan kenyamanan penerapan situs web, database, API aplikasi, dan mesin model pembelajaran (ML). Dari hasil penggunaan docker ini kami mendapati sistem yang lebih efisien dari segi *running time apps*.

Metode klasifikasi, yaitu menggunakan *Machine Learning*. Dengan menggunakan 3 fitur data yang kami ubah dari Kaggle, 3 fitur data ini terdiri dari Blood Pressure, Diabetes PedigreeFunction, dan Outcome untuk hasil deteksinya. Ini menunjukkan jika 3 fitur yang kami pilih sebagai atribut klasifikasi dapat dijadikan rujukan deteksi diabetes.

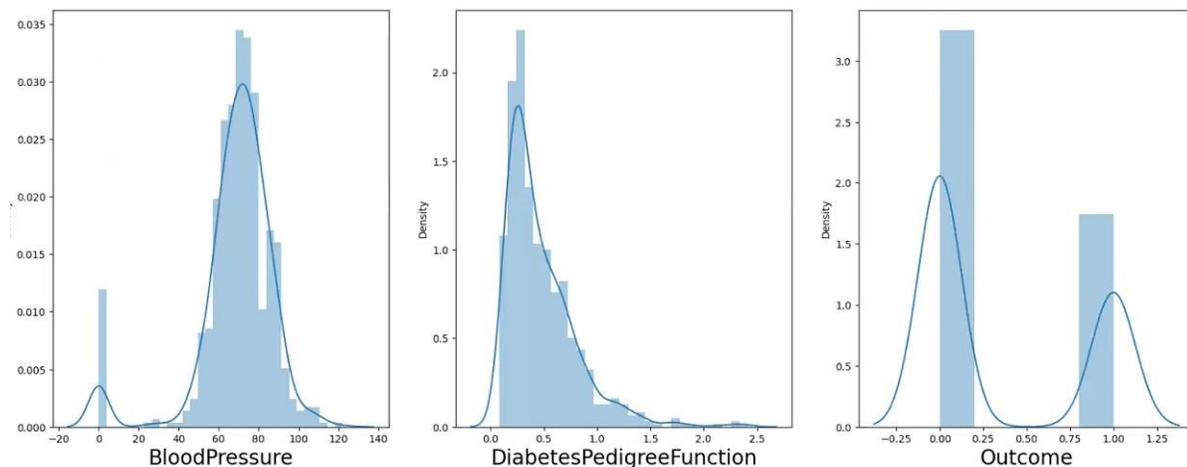
Dengan adanya penelitian ini diharapkan dapat mempermudah pendeteksian penyakit diabetes secara dini sehingga dapat segera ditindak lanjuti.

2. Metode

Metode yang digunakan dalam penelitian ini menggunakan 1 metode klasifikasi, yaitu menggunakan *Machine Learning*. Dengan menggunakan 3 fitur data yang kami ubah dari Kaggle, 3 fitur data ini terdiri dari Blood Pressure, Diabetes Pedigree Function, dan Outcome untuk hasil deteksinya.

Menggabungkan teknologi docker dan machine learning sebagai fitur deteksi diabetes untuk lebih memudahkan dalam implementasi sistem deteksi, docker adalah layanan containerisasi yang memungkinkan kenyamanan penerapan situs web, database, API aplikasi, dan mesin model pembelajaran (ML).

Dalam hal ini *Machine Learning* memiliki kemampuan untuk mempelajari data yang adadan data yang ia peroleh sehingga bisa melakukan tugas tertentu. Tugas yang dapat dilakukan oleh *Machine Learning* pun sangat beragam, tergantung dari apa yang ia pelajari.



Gambar 1. 3 Jenis fitur data.

3. Hasil dan Pembahasan

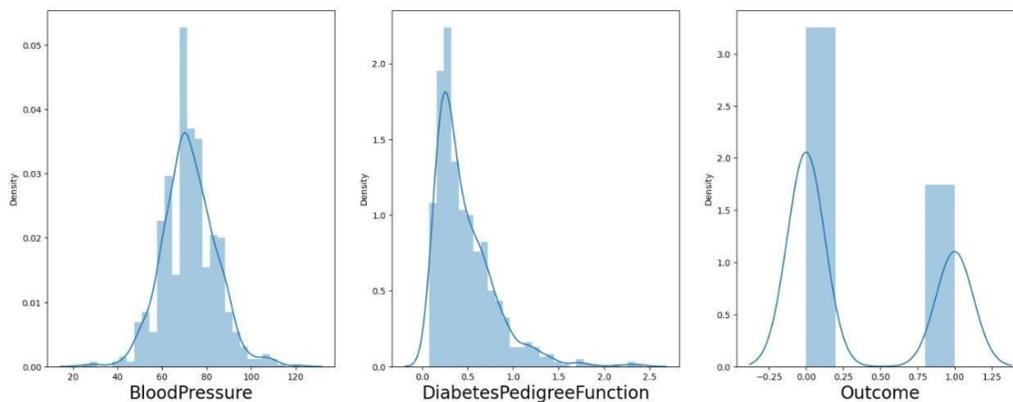
A. Data

Data yang digunakan dalam penelitian ini menggunakan dataset dari Kaggle yang kemudian kita implementasikan dalam Machine Learning yang menggunakan Docker dan Flask pada Python sebagai alat untuk melakukan deteksi Diabetes Melitus.

Teknologi Docker dan Machine Learning dapat menghasilkan akurasi yang tinggi dalam pengolahan dataset survey diabetes yang sudah ada. Dengan menghitung dan mempertimbangkan Blood Pressure serta Pedigree Function sehingga Outcome yang dihasilkan menggunakan perhitungan dari Machine Learning dapat menjadi rujukan untuk system deteksi penyakit Diabetes Melitus.

B. Pengujian

Gambar 2 merupakan hasil perubahan data untuk menghilangkan data yang bernilai null dengan menggunakan metode average



Gambar 2. Hasil perubahan data untuk menghilangkan data yang tidak valid (null)

Dari gambar 3 menunjukkan bahwa data hasil klasifikasi yang ditunjukkan dengan label outcome 1 bernilai diabetes sedangkan 0 tidak. Pada tekanan darah 72 menunjukkan positif mengidap penyakit diabetes melitus sedangkan tekanan darah 66 menunjukkan negative.

	BloodPressure	DiabetesPedigreeFunction	Outcome
0	72	0.627	1
1	66	0.351	0
2	64	0.672	1
3	66	0.167	0
4	40	2.288	1

Gambar 3. Hasil klasifikasi pengidap penyakit Diabetes Melitus

Pada exploratory data analysis dibawah ini menunjukkan bahwa Count merupakan jumlah dari orang yang dilakukan survey diabetes, sehingga masuk kedalam dataset Kaggle, Mean merupakan nilai rata-rata yang diperoleh dari hasil penghitungan sebanyak 768 orang, Std merupakan simpangan terjauh dari penghitungan data ini sendiri yaitu sebesar 12, Blood Pressure minimum sebesar 24, sedangkan 25, 50, dan 75% merupakan presentase orang yang menunjukkan rata-rata tekanan darah dan yang terakhir yaitu data blood pressure maksimal.

	BloodPressure	DiabetesPedigreeFunction	Outcome
count	768.000000	768.000000	768.000000
mean	72.254807	0.471876	0.348958
std	12.115932	0.331329	0.476951
min	24.000000	0.078000	0.000000
25%	64.000000	0.243750	0.000000
50%	72.000000	0.372500	0.000000
75%	80.000000	0.626250	1.000000
max	122.000000	2.420000	1.000000

Gambar 4. Exploratory data analysis

Pada gambar 5 merupakan Confusion matrix sehingga menghasilkan akurasi dari dataset yang diolah menggunakan Machine Learning.

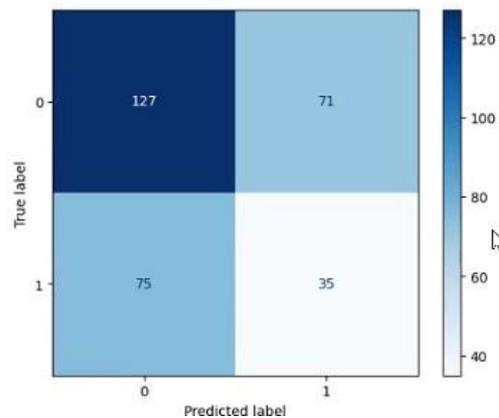
Rumus :

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

Keterangan :

TP = True Positif, TN = True Negatif, FP = False Positif, FN = False Negatif.

Setelah kita hitung hasil akurasinya sebesar 52% yang didapat dari perhitungan diatas menggunakan Machine Learning.



Gambar 5. Confusion Matrix

Penjelasan dari Confusion Matrix :

- 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.

4. Simpulan

Dari hasil klasifikasi kita menggunakan 1 metode klasifikasi *Machine Learning* yang menghasilkan > 50% akurasi dari data training dan testing. Ini menunjukkan jika 2 fitur yang kami pilih sebagai 6tribute klasifikasi dapat dijadikan rujukan deteksi Diabetes Melitus.

Daftar Pustaka

- Mujumdar, A., & Vaidehi, V. (2019). Diabetes Prediction using Machine Learning Algorithms. *Procedia Computer Science*, 165, 292–299. <https://doi.org/10.1016/j.procs.2020.01.047>
- Khanam, J. J., & Foo, S. Y. (2021). A comparison of machine learning algorithms for diabetes prediction. *ICT Express*, 7(4), 432–439. <https://doi.org/10.1016/j.ict.2021.02.004>
- Khairina, N., Susilawati, & Syah, R. (2022). Utilization Docker Swarms in a Container Technology System To Problem Solving Load Balance. *Journal of Theoretical and Applied Information Technology*, 100(18), 5201–5208.
- Openja, M., Majidi, F., Khomh, F., Chembakottu, B., & Li, H. (2022). *Studying the Practices of Deploying Machine Learning Projects on Docker*. *ACM International Conference Proceeding Series* (Vol. 1).

Association for Computing Machinery. <https://doi.org/10.1145/3530019.3530039>

- Dwiyatno, S., Rachmat, E., Sari, A. P., & Gustiawan, O. (2020). Implementasi Virtualisasi Server Berbasis Docker Container. *PROSISKO: Jurnal Pengembangan Riset Dan Observasi Sistem Komputer*, 7(2), 165–175. <https://doi.org/10.30656/prosisko.v7i2.2520>
- Jahiduddin, A., Pramukantoro, E. S., & Bakhtiar, F. A. (2020). Pengembangan Infrastruktur Analisis Data Heart Rate berbasis Microservices menggunakan Kubernetes. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, 4(1), 100–108.
- Megantara, R. A., Alzami, F., Anggi, R., & Puji, D. (2022). Memaksimalkan Utilitas Server Universitas Pada Masa Covid-19, (April), 48–54.
- Kim, B. S., Lee, S. H., Lee, Y. R., Park, Y. H., & Jeong, J. (2022). Design and Implementation of Cloud Docker Application Architecture Based on Machine Learning in Container Management for Smart Manufacturing. *Applied Sciences (Switzerland)*, 12(13), 1–16. <https://doi.org/10.3390/app12136737>
- Mohammadi, G., Pezeshki, F., Mohammadhosseinzadeh Vatanchi, Y., & Moghbeli, F. (2021). Application of Technology in Educating Nursing Students During COVID-19: A Systematic Review. *Frontiers in Health Informatics*, 10(1), 64. <https://doi.org/10.30699/fhi.v10i1.273>
- Zou, Q., Qu, K., Luo, Y., Yin, D., Ju, Y., & Tang, H. (2018). Predicting Diabetes Mellitus With Machine Learning Techniques. *Frontiers in Genetics*, 9(November), 1–10. <https://doi.org/10.3389/fgene.2018.00515>
- Singla, R., Singla, A., Gupta, Y., & Kalra, S. (2019). Artificial intelligence/machine learning in diabetes care. *Indian Journal of Endocrinology and Metabolism*, 23(4), 495–497. https://doi.org/10.4103/ijem.IJEM_228_19
- Llaha, O., & Rista, A. (2021). Prediction and detection of diabetes using machine learning. *CEUR Workshop Proceedings*, 2872, 94–102.
- Fregoso-Aparicio, L., Noguez, J., Montesinos, L., & García-García, J. A. (2021). Machine learning and deep learning predictive models for type 2 diabetes: a systematic review. *Diabetology and Metabolic Syndrome*, 13(1). <https://doi.org/10.1186/s13098-021>