

MENDETEKSI PERILAKU PENYAKIT THALASSEMIA BERDASARKAN REKAM MEDIS MENGGUNAKAN METODE LSTM

Muhammad Fatkhur Rahman, Mohammad Zoqi Sarwani, Muhammad Misdrum

Jurusan Teknik Informatika, Fakultas Teknologi Informasi, Universitas Merdeka Pasuruan

Jl. Ir. H. Juanda No.68 Kota pasuruan, Jawa Timur

fatkhurrahman@student.unmerpas.ac.id

ABSTRAK

Thalassemia merupakan penyakit genetic kronis yang ditandai dengan gangguan sintesis hemoglobin dan memerlukan pemantauan klinis secara berkelanjutan. Data rekam medis yang tersimpan di rumah sakit memiliki potensi besar untuk dimanfaatkan dalam mendeteksi perilaku penyakit secara otomatis. Penelitian ini bertujuan untuk mendeteksi perilaku penyakit thalassemia berdasarkan rekam medis laboratorium menggunakan metode Long Short-Term Memory (LSTM). Dataset yang digunakan terdiri 1.73 data pasien Sembilan parameter laboratorium, yaitu MCV, HbG, MCH, RBC, S, HbA2, HbA, HbF, and Iron. Tahapan penelitian meliputi pengumpulan data, pra-pemrosesan dengan penanganan missing value menggunakan metode mean imputation, perancangan model LSTM, serta evaluasi performa menggunakan metode mean imputation, menggunakan confusion matrix. Hasil perancangan model LSTM, serta evaluasi performa penelitian menunjukkan bahwa model LSTM mampu mempelajari pola kompleks pada data rekam medis dan memberikan performa yang baik dalam mengklasifikasikan kondisi pasien ke dalam kelas normal dan krisis. Penelitian ini diharapkan dapat menjadi pendukung pengambilan Keputusan klinis dalam pemantauan pasien thalassemia

Kata kunci : *Thalassemia, Rekam Medis, LSTM, Deep Learning, Klasifikasi*

1. PENDAHULUAN

Thalassemia adalah salah satu penyakit genetik paling umum di dunia dan juga menjadi salah satu masalah kesehatan secara global yang serius di beberapa negara seperti Mediterania, Asia Tenggara, dan beberapa wilayah yang ada di Afrika. Penyakit ini disebabkan oleh kelainan pada sintesis rantai globin hemoglobin yang berdampak pada kemampuan sel darah membawa oksigen ke seluruh tubuh. Pasien thalassemia umumnya memerlukan pemantauan kondisi kesehatan yang secara rutin melalui pemeriksaan laboratorium. Data rekam medis yang dihasilkan dari pemeriksaan tersebut terus bertambah dan memiliki potensi besar untuk dimanfaatkan sebagai dasar pengambilan Keputusan klinis apabila dianalisis secara tepat.

Permasalahan yang sering dihadapi dalam praktik klinis adalah proses penilaian kondisi pasien yang masih dilakukan secara manual dan bergantung pada interpretasi subjektif tenaga medis, sehingga berpotensi menimbulkan keterlambatan dalam mendeteksi kondisi krisis. Selain itu, pola perubahan parameter laboratorium pada pasien thalassemia bersifat kompleks dan tidak linier, sehingga sulit dianalisis menggunakan metode konvensional. Oleh karena itu, diperlukan pendekatan berbasis kecerdasan buatan yang mampu mempelajari pola data rekam medis secara lebih akurat dan konsisten.

Penelitian ini bertujuan untuk menerapkan metode Long Short-Term Memory (LSTM) dalam rangka mengklasifikasi kondisi pasien thalassemia ke dalam kategori normal dan krisis berdasarkan data rekam medis laboratorium. Penyelesaian permasalahan dilakukan melalui tahapan pra-pemrosesan data, pembentukan sekuens data, perancangan dan pelatihan model LSTM, serta evaluasi performa model

menggunakan metrik klasifikasi dan confusion matrix. Diharapkan penelitian ini dapat menjadi dasar pengembangan sistem pendukung keputusan klinis untuk membantu deteksi dini kondisi krisis pada pasien thalassemia.

2. TINJAUAN PUSTAKA

Dalam penelitian ini, sedikit banyaknya terinspirasi dari referensi penelitian sebelumnya yang relevan dalam konteks dan penelitian sebelumnya menunjukkan bahwa metode LSTM memiliki performa yang baik dalam bidang kesehatan.

[1] Membandingkan metode LSTM dan gated recurrent unit (GRU) dalam prediksi harga emas dan menunjukkan bahwa LSTM mampu menangkap pola deret waktu dengan baik. Penelitian serupa [2] menerapkan LSTM untuk memprediksi harga beras berdasarkan faktor cuaca dan membutuhkan keunggulan LSTM dalam memodelkan hubungan nonlinier pada data series.

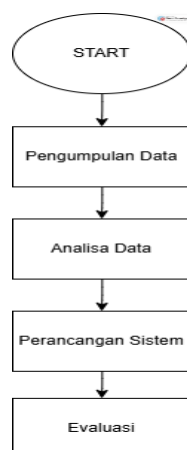
Dalam bidang Kesehatan, beberapa penelitian menyoroti pentingnya pemanfaatan data dan teknologi untuk mendukung pengelolaan penyakit thalassemia. [3] melakukan penyuluhan Kesehatan terkait penyakit thalassemia sebagai upaya peningkatan pengetahuan umum Masyarakat. [4], [5] membahas kualitas hidup anak thalassemia melalui studi tinjauan literatur dan menekankan pentingnya pemantauan kondisi pasien secara berkelanjutan. [6] Melakukan studi retrospektif analisis DNA pasien thalassemia di provinsi Jambi dan memberikan gambaran karakteristik genetic pasien thalassemia di Indonesia. [7] Menunjukkan bahwa kadar hemoglobin pada pasien thalassemia mengalami fluktuasi yang berbeda-beda meskipun telah diberikan transfusi darah, dengan sebagian pasien menunjukkan tren

peningkatan, penurunan, maupun kondisi yang relative stabil. [8] Menunjukkan bahwa Sebagian besar anak usia sekolah penyandang thalassemia berada pada kategori kualitas hidup rentan, terutama pada domain fungsi sekolah.

Penelitian lain [9]menelitian hubungan pengetahuan keluarga mengenai terapi kelasi besi dengan kepatuhan pengobatan pada anak thalassemia. Selain itu, [10]mengulas penerapan aplikasi mobile untuk meningkatkan kuliatas hidup pasien thalassemia. Berdasarkan berbagai penelitian tersebut, dapat disimpulkan bahwa masih diperlukan pendekatan berbasis data rekam medis dan kecerdasan buatan untuk mendeteksi dan memantau perilaku penyakit thalassemia secara objektif. Oleh karena itu, penelitian ini berfokus pada penerapan LSTM untuk klasifikasi kondisi pasien thalassemia berdasarkan data rekam medis laboratorium.

3. METODE PENELITIAN

Penelitian ini akan menggunakan beberapa tahapan dengan diawali pengumpulan data, analisa data, perancangan system, dan yang terakhir evaluasi. Tahapan-tahapan tersebut akan di jelaskan pada gambar 1 di bawah ini dapat memberikan gambaran umum proses penelitian.



Gambar 1. Alur Penelitian Sistem

3.1. Pengumpulan Data

Data set yang digunakan di ambil dari rekam medis laboratorium rumah sakit dengan total 1.073 data pasien. Atribut yang di gunakan meliputi MCV, HbG, MCH, RBC, S, HbA2, HbA, HbF, dan Iron

3.2. Analisa Data

Pada tahap ini pra-pemrosesan, dilakukan idenfikasi data kosong (*missing valaue*). Penanganan *missing value* dilakukan meggunakan metode *mean imputation*, yaitu mengganti nilai kosong dengan nilai rata rata pada kolom yang sesuai

$$x = \frac{\sum x}{n}$$

Keterangan :

- X = data
- n = jumlah data

3.3. Model LSTM

Model LSTM dirancang dengan layer input oleh beberap fitur:

a. Hidden

Salah satu layer LSTM dirancang untuk sebagai lapisan tersembunyi yang bertugas mempelajari dan mengekstraksi pola kompleks dari data rekam medis berbentuk sekuens. Pada penelitian ini digunakan satu hidden layer LSTM dengan 64 neuron yang mampu menangkap ketergantungan jangka pendek maupun jangka panjang pada data laboratorium pasien.

b. Output Data

Output layer terdiri dari satu neuron dengan fungsi aktivasi sigmoid. Lapisan ini menghasilkan nilai probabilitas antara 0 dan1 yang mempresentasikan kondisi pasien. Nilai mendekati 0 menunjukkan kondisi normal, sedangkan nilai 1 menunjukkan kondisi kritis.

c. Fungsi Aktivas

Disini yang digunakam pada hidden layer LSTM Adalah fungsi tanh (*hyperbolic tangent*). Fungsi aktivasi dipilih karena mampu menghasilkan keluaran dalam rentang -1 hingga 1, sehingga membantu menjaga stabilitas gradien selama proses pelatihan dan mempercepat konvergensi model

d. Fungsi loss

Untuk fungsi loss ini mengukur perbedaan antara nilai prediksi probalitas yang dihasilkan model dengan label sebenarnya. Binary Cross-Entropy dirumuskan sebagai berikut:

$$L = -[y \log(p) + (1 - y) \log(1 - p)]$$

Dimana yang merupakan label sebenarnya dan p adalahDenga meminimmalkan nilai loss ini selama proses pelatihan, model LSTM diharapkan mampu meningkatkan akurasi klasifikasi kondisi pasien thalassemia

3.4. Evaluasi

Pada tahap terakhir ini melakukan evaluasi. Dengan ini menggunakan *confusion matrix* untuk mengevaluasi model dari LSTM.

Confusion matrix

Confusion matrix adalah sebuah tabel ringkasan yang digunakan untuk mengukur performa model klasifikasi dengan menunjukkan jumlah benar dan salah untuk setiap kelas.

Akurasi

Mengukur seberapa total banyak jumlah data positif yang benar dari total keseluruhan data.

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

Precision

Mengukur ketepatan model dalam memprediksi suatu kelas.

$$Precision = \frac{TP}{TP + FP}$$

Recall

Mengukur kemampuan model dalam menemukan data yang benar termasuk dalam suatu kelas.

$$\text{Recall} = \frac{TP}{TP + FN}$$

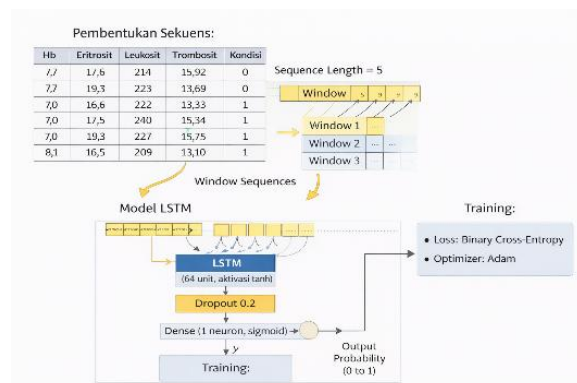
F1-Score

Menggabungkan hasil dari precision dan recall antara katagori hasil prediksi dengan kategori yang sebenar-benarnya.

$$F1 = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}$$

4. HASIL DAN PEMBAHASAN

Pada tahap ini melakukan pengujian model *Long Short-Term Memori* (LSTM) menggunakan data rekam medis yang telah melalui proses pra-promosesan dan pembentukan data sekuens. Model dilatih selama 60 epoch dengan ukuran batch 16 serta menggunakan optimizer Adam dan fungsi *los Binary Cross-Entropy*. Proses pelatihan menghasilkan dua grafik utama, yaitu grafik akurasi dan garafik loss, yang di gunakan untuk menganalisis kinerja model.



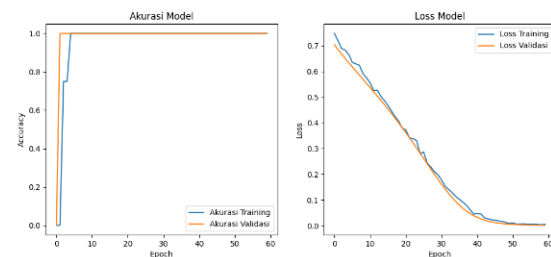
Gambar 2. Arsitektur Implementasi Model LSTM

Gambar di atas menunjukkan tahapan implementasi metode Long Short-Term Memori (LSTM) yang digunakan dalam penelitian ini. Pada tahap awal dilakukan pembentukan sekuens data, di amana data rekam medis laboratorium di ubah menjadi *window time-series* dengan panjang sekuens (sequence length) sebesar 5. Pembentukan sekuens ini bertujuan agar model LSTM mampu mempelajari pola temporal serta ketergantungan antar data secara berurutan.

Selanjutnya, data sekuens tersebut di proses menggunakan model LSTM yang terdiri dari satu hidden layer dengan 64 unit neuron dan fungsi aktivasi *tanh*. Lapisan ini berfungsi untuk mengekstraksi pola kompleks dan hubungan jangka panjang data time-series. Untuk mengurangi resiko *overfitting*, di tambahkan dropout layer dengan nilai 0,2 yang berfungsi menonaktifkan sebagian neuron secara acak selama proses pelatihan

Pada bagian akhir arsitektur, digunakan output layer berupa satu neuron dengan fungsi aktivasi sigmoid yang menghasilkan nilai probabilitas antara 0 dan 1. Nilai probabilitas tersebut di gunakan untuk

menentukan klasifikasi kondisi pasien ke dalam kelas normal atau krisis. Proses pelatihan model dilakukan menggunakan optimizer adam dengan fungsi loss Binary Cross-Entropy, yang sesuai untuk permasalahan klasifikasi biner. Kombinasi arsitektur dan parameter pelatihan ini memukungkan model LSTM mencapai performa klasifikasi yang optimal



Gambar 3. Grafik Akurasi dan Loss Model LSTM

Berdasarkan grafik akurasi, terlihat bahwa akurasi training dan validasi meningkatkan sangat cepat pada beberapa epoch awal dan mencapai nilai mendekati 100%. Hal ini menunjukkan bahwa model LSTM mampu mempelajari pola hubungan antara akurasi training dan validasi juga mengindikasikan bahwa model tidak mengalami overfitting secara signifikan.

Gambar 3 menunjukkan grafik akurasi model LSTM selama proses pelatihan. Terlihat bahwa akurasi data training dan validasi mengalami peningkatan yang signifikan sejak epoch awal dan kemudian stabil hingga akhir pelatihan. Hal ini menandakan bahwa model memiliki kemampuan generalisasi yang baik terhadap data uji.

Grafik loss menunjukkan penurunan nilai loss secara bertahap seiring bertambahnya jumlah epoch, baik pada data training maupun data validasi. Nilai loss yang semakin kecil medekati nol menandakan bahwa kesalahan prediksi model semakin berkurang dan proses pembelajaran berjalan dengan baik. Dengan Gambar 3 kurva loss menurun secara konsisten dan saling berdekatan menunjukkan bahwa proses pelatihan berlangsung stabil tanpa indikasi overfitting.

Grafik loss menunjukkan penurunan nilai loss secara bertahap seiring bertambahnya jumlah epoch, baik pada data training maupun data validasi. Nilai loss yang semakin kecil medekati nol menandakan bahwa kesalahan prediksi model semakin berkurang dan proses pembelajaran berjalan dengan baik. Dengan Gambar 3 kurva loss menurun secara konsisten dan saling berdekatan menunjukkan bahwa proses pelatihan berlangsung stabil tanpa indikasi overfitting.

Model LSTM dibangun dengan satu lapisan LSTM berjumlah 64 unit neuron dan fungsi aktivasi tanh, diikuti oleh lapisan dropout sebesar 0,2 untuk mengurangi risiko overfitting. Lapisan output menggunakan satu neuron dengan fungsi aktivasi sigmoid yang menghasilkan probabilitas kondisi pasien. Nilai (0) dan krisis (1) dengan ambang batas 0,5.

Hasil evaluasi model ditunjukkan pada Tabel 1 yang menyajikan metrik evaluasi berupa akurasi, precision, recall, dan F1-score. Nilai metrik yang tinggi menunjukkan bahwa model LSTM memiliki performa yang sangat baik dalam mengklasifikasikan kondisi pasien thalassemia.

Tabel 1. Hasil Evaluasi Model LSTM

Metrik	Nilai
Akurasi	0,99
Precision	0,99
Recall	0,98
F1-Score	0,99

Akurasi menunjukkan Tingkat ketepatan model dalam memprediksi kondisi pasien secara keseluruhan. Precision mengukur ketepatan model dalam memprediksi kelas krisis, sedangkan recall menunjukkan kemampuan model dalam mendeteksi seluruh data data pasien yang benar-benar berada pada kondisi krisis. F1-score merupakan nilai harmonis antara precision dan recall yang mencerminkan performa klasifikasi secara menyeluruh.

Seluruhnya, hasil pengujian menunjukkan bahwa model LSTM memiliki kemampuan yang kuat dalam mempelajari pola data sekuens pada rekam medis dan berpotensi untuk dikembangkan lebih lanjut sebagai sistem pendukung keputusan klinis dalam pemantauan pasien thalassemia.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil menerapkan metode *long short-term memory* (LSTM) untuk mampu mengklasifikasikan kondisi pasien thalassemia ke dalam kategori normal dan krisis secara efektif berdasarkan data rekam medis laboratorium. Model LSTM memiliki performa yang baik dengan Tingkat sensitivitas dan spesifisitas yang tinggi, sehingga mampu mendeteksi kondisi krisis secara akurat sekaligus mengklasifikasikan kondisi normal dengan tepat. Hasil ini menunjukkan bahwa LSTM berpotensi dikembangkan sebagai sistem pendukung Keputusan klinis untuk membantu deteksi dini dan pemantauan kondisi pasien thalassemia, serta dapat menjadi dasar bagi penelitian lanjutan dengan penambahan data dan parameter klinis yang lebih beragam.

DAFTAR PUSTAKA

- [1] A. Tholib, N. K. Agusmawati, and F. Khoiriyah, "PREDIKSI HARGA EMAS MENGGUNAKAN METODE LSTM DAN GRU," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 11, no. 3, Aug. 2023, doi: 10.23960/jitet.v11i3.3250.
- [2] Y. Ashari and A. Suhendar, "IMPLEMENTASI ALGORITMA LONG SHORT-TERM MEMORY (LSTM) UNTUK MEMPREDIKSI HARGA BERAS DI JAWA TENGAH BERDASARKAN CUACA," *Jurnal Teknologi Informasi*, vol. 5, no. 3, 2024, doi: 10.46576/djtechno.
- [3] I. Aisyahi, D. Setiadi Kelana, S. Diploma, and T. Keperawatan, "Penyuluhan Kesehatan Tentang Penyakit Thalassemia." [Online]. Available: <https://ejournal.upi.edu/index.php/BSM/index>
- [4] L. Fatkuriyah and A. Hidayati, "FACTORS RELATED TO QUALITY OF LIFE AMONG CHILDREN WITH THALASSEMIA MAJOR: A LITERATURE REVIEW," *Nurse and Health: Jurnal Keperawatan*, vol. 11, no. 1, pp. 47–56, Jun. 2022, doi: 10.36720/nhjk.v11i1.340.
- [5] R. Rahmah and S. N. N. Makiyah, "Quality Of Life of Children with Thalassemia in Indonesia: Review," *IJNP (Indonesian Journal of Nursing Practices)*, vol. 6, no. 1, pp. 1–7, Jun. 2022, doi: 10.18196/ijn.v6i1.10477.
- [6] S. Hutabarat and W. I. D. Aurora, "A Retrospective Study of DNA Analysis Thalassemia Patients in Jambi Province for the Years 2016-2020," *Jambi Medical Journal : Jurnal Kedokteran dan Kesehatan*, vol. 12, no. 1, pp. 01–07, May 2024, doi: 10.22437/jmj.v12i1.30868.
- [7] "pengelola,+REZA+24+-+29".
- [8] H. S. Mediani, G. G. Ramdhanie, and A. Fikri, "Kualitas Hidup Anak Usia Sekolah Penyandang Thalassemia," *Jurnal Obsesi : Jurnal Pendidikan Anak Usia Dini*, vol. 6, no. 3, pp. 2243–2250, Jan. 2022, doi: 10.31004/obsesi.v6i3.1933.
- [9] A. Setiyadi, I. Parulian, and M. Ulfan, "RELATIONSHIP BETWEEN FAMILY KNOWLEDGE ABOUT IRON CHELATION AND ADHERENCE TO GIVING IRON CHELATION IN CHILDREN WITH THALASSEMIA AT RSCM KIARA," vol. II, 2022, [Online]. Available: <http://ejournal.fkmumj.ac.id/>
- [10] A. A. Syaiful and S. Mulyono, "Mobile App Deployment for Quality-of-Life Patient with Thalassemia: A narrative review," *Jurnal Berita Ilmu Keperawatan*, vol. 16, no. 1, p. 2023