

Klasifikasi *HIV AIDS* dengan Aplikasi *Rapid Miner*

Aden Wahyu P. Samudra¹, Rizky A. Susanto², Alfian R Putra³, Felix Indra Kurniadi⁴, Budi Juarto⁵

Computer Science Department, School of Computer Science, Universitas Bina Nusantara
aden.samudra@binus.ac.id, rizky.susanto@binus.ac.id, alfian.putra@binus.ac.id, felix.indra@binus.ac.id,
budi.juarto@binus.ac.id

Diterima : 22 Agustus 2022

Disetujui : 01 Oktober 2022

Abstract— HIV adalah virus yang menyerang sistem kekebalan tubuh yang selanjutnya meningkatkan kemampuan tubuh untuk melawan infeksi dan penyakit. Sejarah AIDS Virus HIV dikatakan berasal dari Kinshasa, Republik Demokratik Kongo. Pada saat itu, para ahli percaya bahwa HIV berasal dari spesies simpanse yang ditularkan ke manusia. Pada simpanse, virus tersebut diberi nama Simian Immunodeficiency Virus atau SIV. Sebelum kemudian menyebabkan penularan HIV pada manusia, penularan virus simpanse ini mungkin berasal dari perburuan simpanse untuk diambil dagingnya, kemudian para pemburu tersebut terkena darah hewan yang terinfeksi. Studi oleh Pusat Pencegahan dan Pengendalian Penyakit (CDC) menunjukkan bahwa HIV mungkin telah ditularkan dari simpanse ke manusia sejak akhir 1800-an. Kinshasa adalah kota terbesar di Kongo, kota dengan pertumbuhan tercepat dengan jaringan transportasi yang menjangkau seluruh negeri. Sebuah laporan menyebutkan sejarah di balik penularan HIV AIDS dari Kongo ke seluruh dunia. Maraknya perdagangan seks, pertumbuhan penduduk, dan jarum suntik yang tidak steril di klinik-klinik diduga menjadi penyebab penyebaran virus HIV yang cukup pesat saat itu. Sejarah juga mencatat AIDS kemudian merajalela di Amerika, Eropa, lalu ke seluruh dunia. Untuk memeriksa data yang ada kami menggunakan aplikasi Data Miner untuk memudahkan kami dalam memeriksa data tersebut. masih banyak pasien yang terpapar virus HIV yang artinya masih banyak masyarakat yang tidak sadar akan bahaya virus yang jika tidak segera ditangani virus ini akan memasuki fase akhir yang sangat berbahaya atau yang kita ketahui sebagai AIDS.

Keywords —HIV, machine learning, rapid miner

I. PENDAHULUAN

Human Immunodeficiency Virus (HIV) adalah sebuah virus yang menyerang sistem kekebalan tubuh dengan cara merusak sel darah putih terutama bagian limfosit dalam tubuh. Hal ini mengakibatkan turunnya sistem kekebalan tubuh penderita HIV. Beberapa orang bahkan mampu hidup dengan tidak tampak seperti seseorang yang terkena penyakit HIV dan beberapa kasus sampai tidak membutuhkan pengobatan [1].

Beberapa faktor yang menyebabkan penularan penyakit HIV salah satunya adalah

hubungan badan yang berisiko dan berbagi alat suntik dengan orang lain. Dilain pihak HIV tidak dapat menular dari cairan tubuh seperti saliva dan lainnya [1].

Berdasarkan data dari Kementerian Kesehatan Republik Indonesia yang dikeluarkan pada tahun 2015. Pada Data tersebut menyatakan bahwa pada tahun 2010-2012 jumlah kasus baru HIV positif tidak mengalami perubahan yang signifikan akan tetapi pada tahun 2013 sampai 2014 terdapat kenaikan yang cukup signifikan [2].

Acquired Immune Deficiency Syndrome merupakan gejala dan infeksi yang muncul

dikarenakan kerusakan sistem kekebalan tubuh yang dikarenakan oleh virus HIV. Orang yang terkena ini akan rentan terhadap penyakit infeksi *opportunistic* ataupun tumor[3].

Berdasarkan resiko yang didapatkan dari AIDS maka dibutuhkan sebuah pendekatan machine learning yang menggunakan pendekatan Data mining. Beberapa penelitian yang menggunakan pendekatan data mining dalam menyelesaikan problematika mengenai HIV/AIDS salah satunya adalah yang dilakukan oleh Noor et al., [4] pada penelitian ini menggunakan metode K-Means clustering. Hasil dari penelitian ini menemukan bahwa k-means memberikan 3 cluster. Pendekatan menggunakan metode k-means lainnya dilakukan oleh [5] hasil yang diberikan juga memenuhi kriteria yang diharapkan oleh penulis.

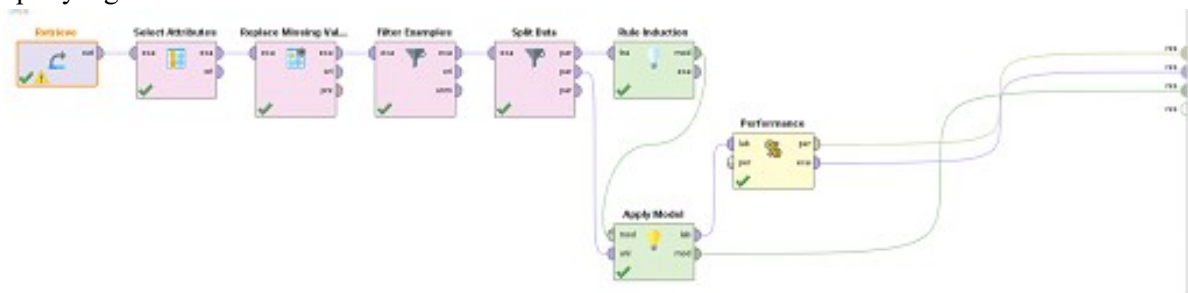
Pada artikel ilmiah ini akan dibagi menjadi empat bab yang akan menjelaskan secara rinci apa yang dilakukan. Pada Bab II kami akan

menjelaskan mengenai metode yang dilakukan dalam penelitian, Bab III hasil dari penelitian dan pembahasan singkat. Bab IV merupakan Kesimpulan.

II. METODE

Pertama kami menggunakan data HIV/AIDS yang disediakan oleh MoleculeNet. Data ini terdiri atas fitur senyawa kimia Simplified Molecular Input Line Entry Specification (SMILES) [6]. Proses selanjutnya kami menggunakan library RDKit untuk mengekstraksi fitur dari SMILES yang merubahnya menjadi sebuah fitur vector yang nanti nya akan diolah dalam rapid miner.

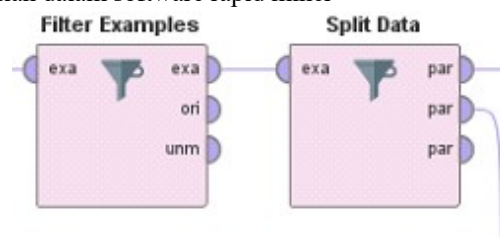
Pada tahapan setelah melakukan proses fitur ekstraksi kami menggunakan software rapid miner untuk mengolah data yang kami miliki. Gambar 1 merupakan rangkaian tahapan yang kami lakukan.



Gambar 1. Rangkaian metodologi penelitian dalam software rapid miner

Tahapan pertama adalah dimulai dengan memasukan operator retrieve untuk mengimport data. Selanjutnya dilakukan proses pemilihan attribute yang akan menjadi target dan fitur. Selanjutnya akan dilakukan proses uji normalitas data dengan menghilangkan data yang memiliki missing value atau menggunakan proses replace missing.

Tahapan berikutnya adalah melakukan splitting data menjadi 70: 30 dan 80:20.



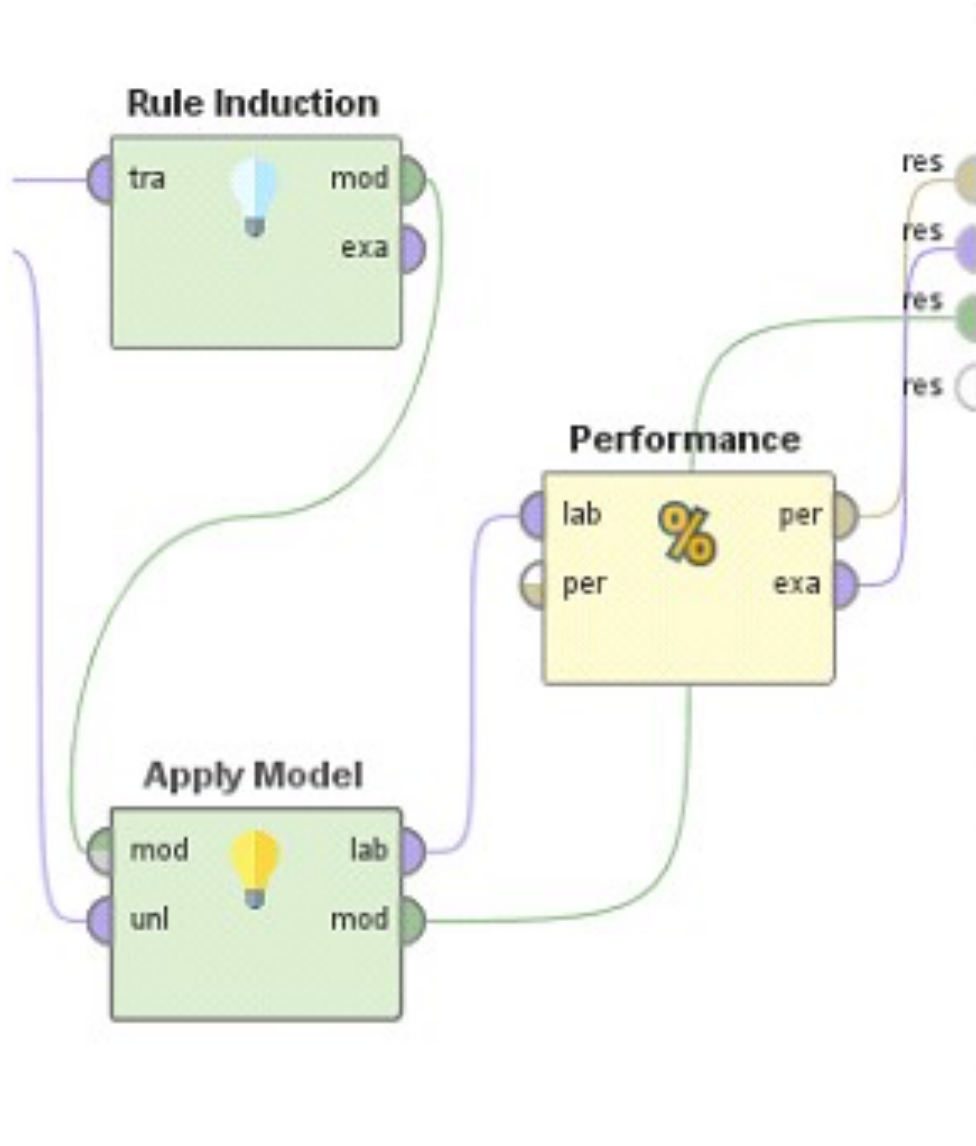
Gambar 2. Proses Splitting Data

Selanjutnya dilakukan proses penerapan dengan menggunakan beberapa penerapan yang dapat dijelaskan sebagai berikut:

- Rule induction berfungsi untuk mengatur model penelitian yang akan dilakukan dengan cara menghubungkan mod point ke mod apply model point. Mod di sini adalah modelnya.

- Menerapkan fungsi model untuk melakukan perhitungan berdasarkan model yang telah diatur dalam aturan induksi. Setiap titik fungsi dalam model yang diterapkan harus terhubung dan digabungkan dengan operator lain. mod terhubung ke induksi aturan, uni terhubung ke data split, lab terhubung ke kinerja dan mod terhubung ke titik akhir.
- Performance berfungsi untuk menampilkan hasil penerapan model yang telah ditentukan pada rule induction dan apply model.

Gambar 3 akan menjelaskan mengenai apa yang dilakukan dalam rapid miner untuk melakukan proses modelling dan rule induction.



Gambar 3. Rule Induction dan Apply model dalam rapid miner.

Selanjutnya kita mencoba mencari hasil dari modeling menggunakan rapid miner. Berikut

adalah hasil dari menggunakan modeling dengan rapid miner.

Row No.	HIV_active	prediction(H...	confidence(0)	confidence(1)	activity
1	0	0	1	0	CI
2	0	0	1	0	CI
3	0	0	1	0	CI
4	0	0	1	0	CI
5	0	0	1	0	CI
6	0	0	1	0	CI
7	0	0	1	0	CI
8	0	0	1	0	CI
9	0	0	1	0	CI
10	0	0	1	0	CI
11	0	0	1	0	CI
12	0	0	1	0	CI
13	0	0	1	0	CI
14	0	0	1	0	CI
15	0	0	1	0	CI

ExampleSet (12,338 examples, 4 special attributes, 1 regular attribute)

Gambar 4. Hasil prediksi

Berdasarkan uraian di atas, dapat diketahui bahwa data penelitian sebelumnya akan diseleksi dan diganti pada proses normalisasi data terlebih dahulu. Namun, hasilnya ditemukan ada 3.290 data dalam kegiatan tersebut. Selanjutnya kita melakukan filter contoh untuk menghilangkan data yang hilang. Kemudian dalam menganalisa kita membagi data terlebih dahulu dengan perbandingan 70% dan 30%. Kemudian setelah itu akan dilakukan klasifikasi data untuk disimpulkan. Dimana hasil interpretasi menemukan bahwa:

Jika CI = 0, maka CM dan CA positif. Karena aktivitas CI mengakibatkan HIV Negatif. Dalam penelitian ini, HIV Activity disimbolkan sebagai berikut: 0 = Negatif 1 = Positif.

III. HASIL DAN DISKUS

HIV/AIDS mempengaruhi pembangunan manusia di negara-negara Afrika melalui dampak yang menghancurkan pada indikator kesehatan dan demografi seperti harapan hidup saat lahir, bantuan perawatan kesehatan, distribusi usia dan jenis kelamin, indikator ekonomi seperti pendapatan,

angkatan kerja, dan pertumbuhan ekonomi, pendidikan dan pendapatan. . pengetahuan dan indikator lainnya seperti tata kelola, ketidaksetaraan gender dan hak asasi manusia.

Oleh karena itu kami melakukan penelitian tentang HIV/AIDS dengan kumpulan data. Kemudian kami mengolahnya sehingga kami mendapatkan visualisasi dari kumpulan data tersebut. Dan kemudian kami menginterpretasikan visualisasi kumpulan data. Dari hasil penelitian terhadap data-data yang kami olah, dapat kami simpulkan bahwa penelitian kami harus diinvestigasi lebih lanjut dan diolah oleh sumber daya manusia yang berkualitas. Agar tercipta suatu penelitian yang memiliki hasil yang lebih akurat.

IV. KESIMPULAN

Kesimpulan yang dapat diambil adalah masih banyak pasien yang terpapar virus HIV yang artinya masih banyak masyarakat yang tidak sadar akan bahaya virus yang jika tidak segera ditangani virus ini akan masuk ke fase akhir yang sangat berbahaya atau yang lebih kita khawatirkan. akrab dengan AIDS. masih terdapat beberapa kendala, sehingga datanya belum

sepenuhnya akurat. Namun dapat kita simpulkan bahwa penyakit ini bisa menular namun tidak secepat virus pada umumnya. Hal ini dikarenakan cara penularan virus ini tidak semudah virus lainnya, namun harus melalui suatu cara.

DAFTAR PUSTAKA

- [1] Y. T. Gunawan, I. Prasetyowati, and M. Ririanty, "Hubungan Karakteristik ODHA Dengan Kejadian Loss To Follow Up Terapi ARV Di Kabupaten Jember," *J. IKESMA*, vol. 12, no. 1, pp. 53–64.
- [2] Kementerian Kesehatan RI, "Laporan HIV/AIDS Triwulan 1 Tahun 2017," Kementerian Kesehatan Republik Indonesia, Jakarta: Depkes RI. [Online]. Available: http://siha.depkes.go.id/portal/files_upload/Laporan_HIV_AIDS_TW_4_Tahun_2017_1.pdf
- [3] M. R. Prayuda, "Pencegahan dan Tatalaksana HIV/AIDS," *J. Agromed Unila*, vol. 2, no. 3, p. 5.
- [4] H. Noor, A. Dharmawati, and T. W. Qur'ana, "PENERAPAN ALGORITMA K-MEANS CLUSTERING ANALYSIS PADA KASUS PENDERITA HIV/AIDS (STUDI KASUS KABUPATEN BANJAR)," *Technol. J. Ilm.*, vol. 12, no. 2, p. 72, Apr. 2021, doi: 10.31602/tji.v12i2.4573.
- [5] R. I. L. Sinaga, W. Saputra, and H. Qurniawan, "Pengelompokan Jumlah Kasus Penyakit Aids Berdasarkan Provinsi Menggunakan Metode K-Means," vol. 2, no. 2, p. 9, 2021.
- [6] Z. Wu *et al.*, "MoleculeNet: a benchmark for molecular machine learning," *Chem. Sci.*, vol. 9, no. 2, pp. 513–530, 2018, doi: 10.1039/C7SC02664A.