

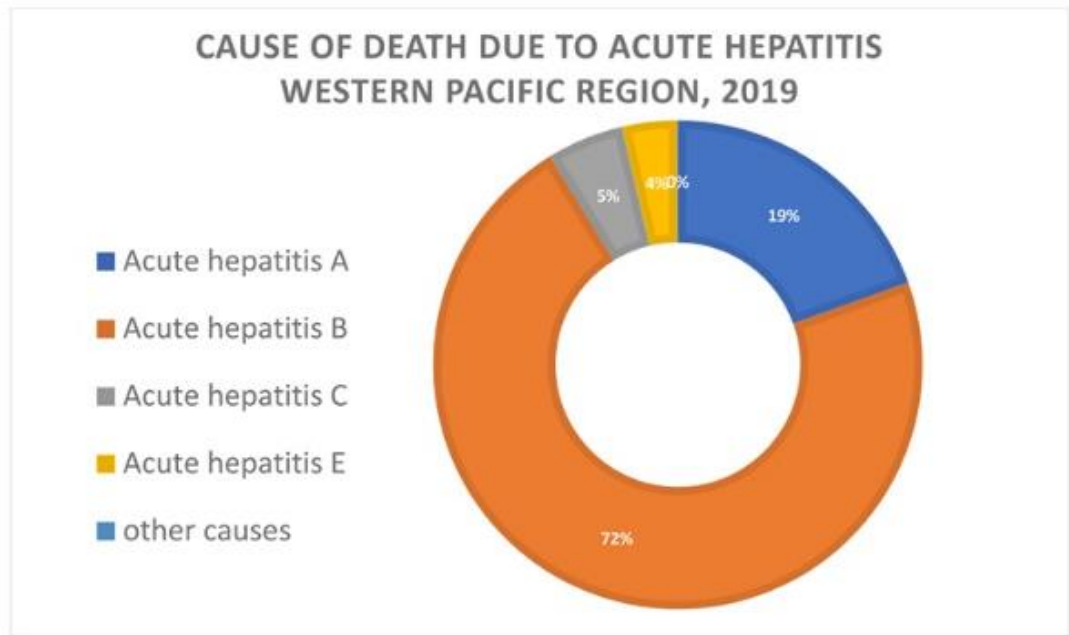
BAB I

PENDAHULUAN

1.1 Latar Belakang

Hepatitis merupakan peradangan pada hati atau liver. Kondisi ini bisa disebabkan oleh beberapa hal, antara lain infeksi virus, kebiasaan minum, minum obat tertentu, penyakit autoimun, dan infeksi cacing hati. Ada beberapa jenis virus hepatitis, seperti: hepatitis a (HAV), hepatitis b (HBV), hepatitis c (HCV), hepatitis d (HDV), dan hepatitis e (HEV)[1]. Setiap jenis hepatitis memiliki penyebaran yang berbeda, seperti hepatitis a umumnya menyebar melalui makanan atau air yang terkontaminasi dan biasanya menyebabkan infeksi singkat tanpa dampak jangka panjang pada hati. Hepatitis B ditularkan melalui kontak dengan cairan tubuh terinfeksi seperti darah atau air mani, dapat mengakibatkan infeksi akut atau kronis yang bisa menyebabkan sirosis hati atau kanker hati. Hepatitis C biasanya menular melalui paparan terhadap darah yang terinfeksi, seringkali terjadi melalui berbagi jarum suntik atau alat obat lainnya. Hepatitis D merupakan jenis hepatitis yang khususnya terjadi pada individu yang telah terinfeksi hepatitis B, menyebar melalui kontak dengan darah terinfeksi atau melalui hubungan seksual. Hepatitis E terutama menular melalui konsumsi makanan atau air yang tercemar, terutama di wilayah dengan sanitasi yang buruk[1].

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A



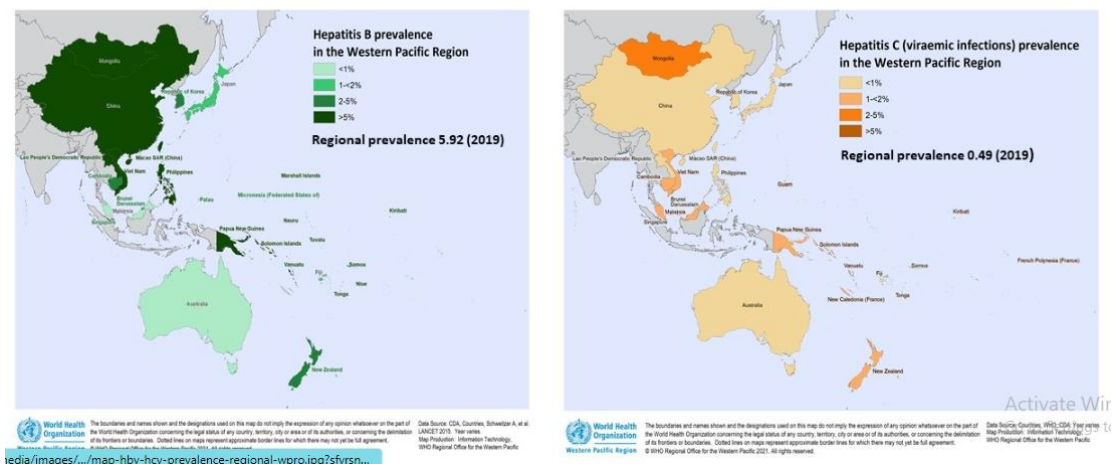
Gambar 1. 1 Grafik Kasus Kematian Hepatitis Akut Tahun 2019

Sumber: [World Health Organization]

Pada Gambar 1.1 merupakan grafik *pie chart* yang dilakukan oleh Organisasi Kesehatan Dunia. Menurut data dari *World Health Organization* (WHO), daerah Pasifik Barat memiliki tingkat kejadian tertinggi untuk jenis hepatitis tertentu pada tahun 2019. Hepatitis B menonjol sebagai penyakit yang paling umum di wilayah ini, menyumbang sekitar 72% dari semua kasus hepatitis yang dilaporkan. Diikuti oleh hepatitis A, yang mencatat sekitar 19% dari total kasus. Hepatitis C juga merupakan masalah serius di daerah ini, meskipun dengan tingkat kejadian yang lebih rendah, yaitu sekitar 5%. Sementara itu, hepatitis E, meskipun lebih jarang dilaporkan, masih menjadi perhatian dengan sekitar 4% dari semua kasus hepatitis di daerah Pasifik Barat[1].

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

Hepatitis B and C prevalence in the Western Pacific Region

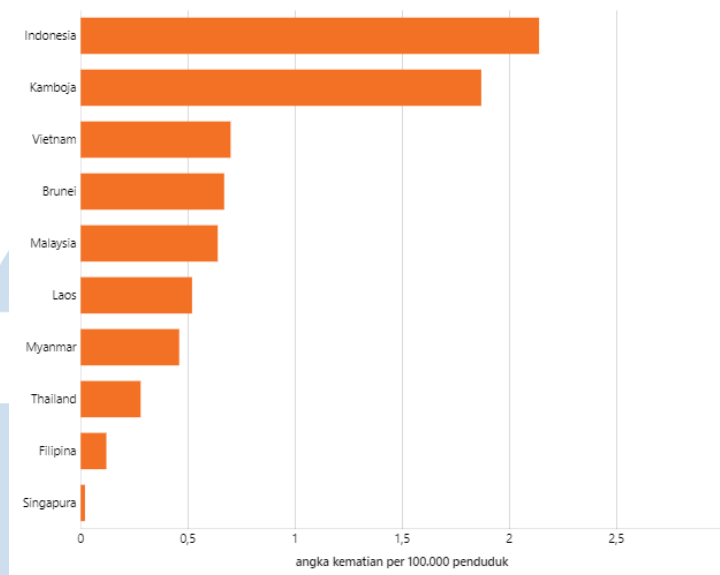


Gambar 1. 2 Grafik Polulasi Hepatitis B dan C di Pasifik Barat

Sumber: [World Health Organization]

Pada Gambar 1.2 merupakan populasi hepatitis B dan C di pasifik barat. Menurut *World Health Organization* (WHO) prevalensi hepatitis B pada masyarakat umum di wilayah tersebut pada tahun 2019 adalah sebesar 5,92%, sementara prevalensi pada anak di bawah usia 5 tahun mencapai 0,46%. Meskipun demikian, sebanyak 21 dari 37 negara dan wilayah dalam wilayah tersebut telah berhasil mencapai target regional yang ditetapkan, yaitu prevalensi hepatitis B pada anak di bawah usia 5 tahun kurang dari 1%. Prevalensi hepatitis C juga menjadi perhatian, dengan angka 0,49% di antara populasi umum pada tahun yang sama[1].

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A



Gambar 1. 3 Grafik Kematian Hepatitis di ASEAN

Sumber: databoks.katadata.co.id

Pada Gambar 1.3 merupakan grafik tingkat kematian terbanyak akibat terpapar hepatitis akut di ASEAN. Pada Gambar 1. 3 indonesia menempati posisi pertama dalam kasus kematian hepatitis akut tertinggi di dunia sebesar 2,14 per 100.000 penduduk, disusul negara kamboja sebesar 1,87 per 100.000 penduduk, dan disusul negara vietnam sebesar 0,7 per 100.000 penduduk[2]. Menurut *kemkes.go.id* di Indonesia, diperkirakan sekitar 20 juta orang menderita hepatitis, dengan hepatitis B memiliki prevalensi tertinggi. Menurut *CDA Foundation*, angka kematian akibat hepatitis B mencapai 51.100 setiap tahun, sedangkan kematian akibat hepatitis C mencapai 5.942 pada tahun 2016. BPJS Kesehatan melaporkan bahwa pada tahun 2022, sebanyak 2.159 orang meninggal karena sirosis dan kanker hati, yang merupakan dampak dari hepatitis kronis, terutama pada orang yang mengalami hepatitis B pada stadium lanjut[3].

Kementerian Kesehatan Indonesia, melalui juru bicaranya dr. Mohammad Syahril, telah mengambil beberapa langkah mitigasi untuk mencegah penyebaran penyakit Hepatitis Akut. Kementerian Kesehatan Indonesia telah melakukan pengumpulan informasi global dengan cepat, meningkatkan kesadaran masyarakat melalui sosialisasi dan edukasi, memperkuat deteksi

melalui teknologi, serta menyusun pedoman tata laksana. Kesadaran masyarakat dalam menerapkan langkah pencegahan juga ditekankan untuk terhindar dari paparan penyakit hepatitis[4]. Program-program yang dilakukan mencegah dan mengobati pasien terpapar hepatitis harus didukung dengan pemanfaatan teknologi yang selalu berkembang, salah satu pemanfaatan teknologi algoritma *machine learning*. Dengan pemanfaatan algoritma *machine learning* diharapkan rumah sakit atau instansi kesehatan dapat membantu pihak terkait dalam memprediksi penyakit hepatitis.

Pemanfaatan teknologi *machine learning* pernah dilakukan oleh beberapa peneliti sebelumnya. Pada penelitian terdahulu yaitu artikel jurnal yang berjudul “*An Ensemble Learning Approach for Enhanced Classification of Patients With Hepatitis and Cirrhosis*” tahun 2021 telah melakukan analisa pada dataset hepatitis C, dimana artikel jurnal tersebut meneliti tingkat klasifikasi pasien hepatitis dan sirosis pada dataset hepatitis C. Peneliti menggunakan tiga algoritma yaitu *random forest*, *decision tree*, dan *linear regression*. Peneliti juga menggunakan dua metode yaitu *two-features RF* dan *AST/ALT ratio*. Dari hasil yang diberikan berdasarkan akurasi ketiga algoritma tersebut, algoritma *random forest* lebih tinggi dibandingkan *decision* dan *liner regression* yaitu 97%. Jika berdasarkan dua metode tersebut akurasi yang diberikan metode *two-features* sedikit lebih unggul dalam prediksi sirosis yaitu 65% dan *AST/ALT* yaitu 56%[5]. Hasil ini membuktikan bahwa *random forest* terbukti baik dalam klasifikasi dengan akurasi yang diberikan unggul dibandingkan algoritma lain[5].

Penggunaan algoritma klasifikasi hibrida dapat menjadi salah satu alternatif dalam melakukan *modelling*, seperti penelitian yang menggunakan algoritma klasifikasi hibrida untuk memprediksi hepatitis C. Artikel jurnal yang menggunakan hibrida berjudul “*Klasifikasi Hepatitis C virus menggunakan Algoritma C4.5*” tahun 2022. Metode yang digunakan pada penelitian ini adalah dengan membandingkan Algoritma C4.5 sebelum menggabungkan metode AdaBoost dan sesudah penerapan metode AdaBoost. Dataset yang digunakan penelitian ini yaitu dataset hepatitis C. Hasil penelitian menunjukkan

bahwa algoritma *C4.5* sesudah penerapan metode *AdaBoost* memiliki performa yang lebih baik daripada sebelum menggunakan metode *AdaBoost*[6]. Penelitian lain juga mengembangkan algoritma *machine learning* yang digabungkan. Penelitian menggunakan algoritma *random forest* dan *support vector machine* yang digabungkan. Hasil penelitian menunjukan bahwa *penggabungan random forest* dan *support vector machine* memberikan performa yang lebih baik[7].

Pada penelitian ini akan dilakukan perbandingan penggunaan algoritma klasifikasi dengan membandingkan algoritma tunggal *random forest* dan hibrida *C4.5-AdaBoost*. Pemodelan diharapkan dapat membantu pihak rumah sakit dalam memprediksi penyakit hepatitis menjadi lebih mudah dan cepat. Penelitian ini berbeda dengan penelitian yang telah dilakukan sebelumnya, dengan alasan pada penelitian ini dilakukan dengan dataset yang diambil langsung dari Rumah Sakit Umum Daerah (RSUD), serta akan membandingkan algoritma *Random Forest*, *C4.5*, *AdaBoost* dan hibrida *C4.5-AdaBoost* tanpa Optimalisasi *hyperparameter*, dan algoritma *Random Forest*, *C4.5*, *AdaBoost* dan hibrida *C4.5-AdaBoost* menggunakan *hyperparameter*. Dengan hasil yang memiliki tingkat akurasi terbaik kemudian akan menjadi algoritma yang akan digunakan untuk melakukan prediksi.

1.2 Rumusan Masalah

Berikut ini adalah rumusan masalah pada penelitian ini :

1. Bagaimana hasil pembangunan model algoritma *Random Forest*, *C4.5*, *AdaBoost*, dan hibrida *C4.5-AdaBoost* sebelum dan sesudah penggunaan optimalisasi *hyperparameter* dalam menganalisis penyakit hepatitis?
2. Bagaimana hasil perbandingan akurasi, recall, f1-score, dan presisi antara algoritma *Random Forest*, *C4.5*, *AdaBoost*, dan hibrida *C4.5-AdaBoost* sebelum dan sesudah penggunaan optimalisasi *hyperparameter* dalam menganalisis penyakit hepatitis?
3. Bagaimana tingkat presentase dari kurva ROC untuk mendukung hasil perbandingan tiap pemodelan algoritma?

1.3 Batasan Masalah

Terdapat batasan masalah pada penelitian kali ini, berikut adalah batasan masalah penelitian ini:

1. Penelitian ini hanya membahas mengenai prediksi klasifikasi hepatitis menggunakan *Knowledge Discovery in Databases* (KDD) sebagai metodologi penelitian.
2. Hasil perbandingan berdasarkan tingkat akurasi, presisi, recall dan f1-score pada algorithm *Random Forest*, *C4.5*, *AdaBoost*, dan algorithm hibrida *C4.5-AdaBoost*, serta didukung dengan tingkat presentasi ROC.
3. Dataset yang digunakan berasal dari Rumah Sakit Umum Daerah Kota Tangerang yang berfokus pada penyakit Hepatitis yang mencakup semua jenis hepatitis.

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

Adapun tujuan dari dilakukannya penelitian ini adalah sebagai berikut:

1. Membangun model algoritma *Random Forest*, *C4.5*, *AdaBoost*, dan hibrida *C4.5-AdaBoost* sebelum dan sesudah penggunaan optimalisasi *hyperparameter* dalam menganalisis penyakit hepatitis.
2. Mengevaluasi tingkat akurasi, presisi, recall, dan f1-score pada algoritma *Random Forest*, *C4.5*, *AdaBoost*, dan hibrida *C4.5-AdaBoost* dalam analisis penyakit hepatitis.
3. Membandingkan kinerja dan mengetahui model algoritma yang cocok antara algoritma *Random Forest*, *C4.5*, *AdaBoost* dan *C4.5-AdaBoost* sebelum dan sesudah penggunaan Optimalisasi *hyperparameter* berdasarkan akurasi, recall, f1-score, dan presisi dalam memprediksi penyakit hepatitis.

1.4.2 Manfaat Penelitian

Berikut adalah beberapa harapan manfaat dalam penelitian ini yaitu:

1. Mengetahui tingkat keberhasilan algoritma *Random Forest* dan hibrida

C4.5- AdaBoost dalam menganalisis virus hepatitis.

2. Menambah wawasan mengenai menggunakan algoritma *Random Forest* dan *C4.5-AdaBoost* terbaik yang dapat digunakan sebagai acuan untuk penelitian terkait terutama dalam bidang kesehatan terutama hepatitis.
3. Penelitian ini diharapkan dapat meningkatkan kesadaran masyarakat tentang penyakit hepatitis, dan mendorong masyarakat untuk melakukan tes hepatitis.

1.5 Sistematika Penulisan

Pada penilitian ini, sistematika pada penulisan laporan sebagai berikut

- BAB 1 PENDAHULUAN
 - Memabahas mengenai latar belakang, rumusan masalah, batasan masalah, tujuan dan manfaat.
- BAB 2 LANDASAN TEORI
 - Membahas mengenai studi literatur seperti teori, *framework*, tools, algoritma dan menjadi landasan dalam melakukan penelitian ini.
- BAB 3 METODOLOGI PENELITIAN
 - Menjelaskan struktur keseluruhan metodologi dan semua langkah yang diambil untuk menghasilkan solusi penelitian pada permasalahan penelitian ini.
- BAB 4 ANALISIS DAN HASIL PENELITIAN
 - Menjelaskan analisis dan hasil penelitian terdapat pembahasan mengenai hasil pemodelan yang telah dibuat dalam penelitian beserta penjelasan analisisnya.
- BAB 5 SIMPULAN DAN SARAN
 - Menjelaskan pembahasan mengenai kesimpulan serta saran berdasarkan penelitian yang telah dilakukan.