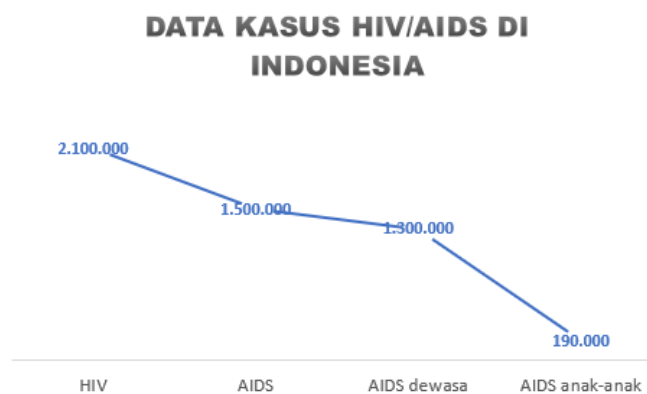


BAB I

PENDAHULUAN

1.1. Latar Belakang

Semakin berkembangnya zaman, data memiliki berbagai macam jenis tergantung bidang apa yang dibutuhkan tidak terkecuali salah satunya adalah bioinformatika. Bioinformatika adalah ilmu yang mempelajari penerapan teknik komputasional untuk mengelola dan menganalisis informasi biologi. Bidang ini mencakup penerapan metode-metode matematika, statistika, dan informatika untuk memecahkan masalah-masalah biologis, terutama dengan menggunakan sekuens DNA dan asam amino serta informasi yang berkaitan dengannya. (Pratiwi, 2018).



Gambar 1.1 Kasus HIV/AIDS di Indonesia

Sumber : (Ditjen PP& PL RI, 2016)

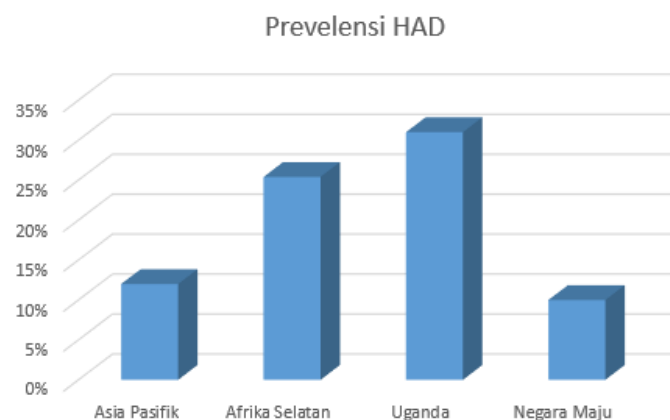
Seperti gambar diatas yakni data pada tahun 2013, jumlah infeksi baru HIV mencapai 2,1 juta dan jumlah kematian akibat AIDS sebanyak 1,5 juta yang terdiri dari 1,3 juta dewasa dan 190.000 anak berusia <15 tahun (Direktorat Jenderal Pengendalian Penyakit dan Penyehatan Lingkungan RI, 2016).

Human Immunodeficiency Virus atau HIV adalah virus yang menyerang sel darah putih di dalam tubuh (limfosit) yang mengakibatkan turunnya kekebalan tubuh manusia. Orang yang dalam darahnya terdapat virus HIV dapat tampak sehat dan belum tentu membutuhkan pengobatan. Meskipun demikian, orang tersebut dapat menularkan virusnya kepada orang lain bila melakukan hubungan

seks berisiko dan berbagi penggunaan alat suntik dengan orang lain. (KPAD Kab. Jember, 2015).

Acquired Immune Deficiency Syndrome atau AIDS adalah sekumpulan gejala penyakit yang timbul karena kekebalan tubuh yang menurun yang disebabkan oleh infeksi HIV. Akibat menurunnya kekebalan tubuh pada seseorang maka orang tersebut sangat mudah terkena penyakit seperti TBC, kandidiasis, berbagai radang pada kulit, paru, saluran pencernaan, otak dan kanker. (KPAD Kab. Jember, 2015).

Epidemiologi demensia (penurunan kemampuan fungsi otak) terkait infeksi yang ter-dokumentasikan dengan baik saat ini adalah demensia terkait HIV-1 (*HIV-1 associated dementia/HAD*).



Gambar 1.2 Prevelensi HAD (HIV-1 Associated Dementia)

Sumber : (Chan dkk, 2012; Robbins dkk, 2011; Ghafouri dkk, 2006)

Terkait gambar diatas pada prevalensi HAD di Asia Pasifik sebesar 12%, di Afrika Selatan dan Uganda masing-masing sebesar 25,4% dan 31%, dan di negara maju sebesar 10%. (Chan dkk, 2012; Robbins dkk, 2011; Ghafouri dkk, 2006). Rendahnya prevalensi HAD di negara maju tersebut berkaitan dengan meningkatnya penggunaan *antiretroviral treatment* (ART), namun penurunan angka kejadian HAD tersebut digantikan oleh meningkatnya angka kejadian gangguan kognitif yang lebih ringan terkait dengan infeksi HIV-1.(Duarte dkk, 2006).

Virus HIV tidak menyebabkan kematian secara langsung pada penderitanya, akan tetapi adanya penurunan imunitas tubuh yang mengakibatkan

mudah terserangnya infeksi oportunistik bagi penderitanya (Fauci & Lane, 2012). Penyakit HIV yang semula bersifat akut dan mematikan berubah menjadi penyakit kronis yang bisa dikelola. Namun demikian, hidup dengan penyakit kronis menyisakan persoalan-persoalan lain yang memerlukan penyesuaian-penyesuaian baik secara fisik, psikologis, sosial, dan spiritual (Lindayani & Maryam, 2017).

Klasifikasi merupakan salah teknik yang sangat populer di dalam data *mining* atau *machine learning*. Klasifikasi bertujuan untuk menemukan suatu fungsi keputusan yang secara akurat dapat memprediksi kelas dari data *testing* dan data *training*. Salah satu metode klasifikasi yang terbaik adalah *Support Vector Machine* (Huang dkk, 2014). *Support Vector Machine* (SVM) adalah suatu pemodelan klasifikasi dengan konsep yang lebih matang dan lebih jelas secara matematis dibandingkan dengan teknik-teknik klasifikasi lainnya. SVM digunakan untuk mencari *hyperplane* terbaik dengan memaksimalkan jarak antar kelas. Hal ini membuktikan bahwa SVM dapat menjamin kemampuan generalisasi yang tinggi untuk data-data yang akan datang (Suyanto, 2017).

Penelitian ini mempunyai misi untuk mengklasifikasikan pasien yang terkena dan tidak terkena HIV yang didasarkan data terunggah dalam situs halaman NCBI oleh Alejandra Borjabad dengan kode GSE28160 yang berjudul “*Significant Effects of Antiretroviral Therapy on Global Gene Expression in Brain Tissues of Patients with HIV-Associated Neurocognitive Disorders*”.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, maka perlu merumuskan masalah yang harus diselesaikan pada penelitian ini, yakni :

1. Bagaimana cara mengolah data ekspresi gen pada penderita HIV-1?
2. Bagaimana hasil klasifikasi data pasien yang terkena HIV-1 dan tidak terkena HIV-1 dengan metode SVM?

1.3. Batasan Masalah

Dalam penelitian ini terdapat batasan masalah supaya penelitian yang dibahas tidak meluas yakni :

1. Aplikasi yang digunakan adalah R 3.5.2 dengan beberapa pilihan *package* bioinformatics yang dibutuhkan.
2. Metode yang digunakan ialah *Support Vector Machine*.
3. Data yang digunakan dari NCBI, yaitu GSE28160.

1.4. Tujuan Penelitian

Beberapa hal yang ingin dituju oleh penulis pada penelitian ini yakni :

1. Mengetahui hasil pengolahan data ekspresi gen penderita HIV-1.
2. Mengetahui hasil klasifikasi pasien HIV-1 dengan metode SVM.

1.5. Manfaat Penelitian

Dari penelitian yang telah dibuat diharapkan dapat memberikan informasi terkait penerapan *machine learning* (SVM) dalam bidang kesehatan pada pasien yang terkena dan tidak terkena HIV, sehingga dapat digunakan sebagai acuan di berbagai kalangan, terkhusus kalangan ilmu kesehatan yang didalamnya terkandung banyaknya pasien yang terkena dan tidak terkena HIV, informatika yang dalam penelitian ini terkandung algoritma-algoritma pemrograman dan juga statistisi yang dalam penelitian ini terkandung perhitungan, langkah-langkah dan juga metode yang terkait dengan statistika.