

BAB I

PENDAHULUAN

1.1. Latar Belakang

Hepatitis adalah kelainan hati berupa peradangan (sel) hati. Peradangan ini ditandai dengan meningkatnya kadar enzim hati. Peningkatan ini disebabkan adanya gangguan atau kerusakan membran hati. Ada dua faktor penyebabnya yaitu faktor infeksi dan faktor non infeksi. Faktor penyebab infeksi antara lain virus hepatitis dan bakteri. Selain karena virus Hepatitis A, B, C, D, E dan G masih banyak virus lain yang berpotensi menyebabkan hepatitis misalnya adenoviruses, CMV, Herpes *simplex* , HIV , rubella ,varicella dan lain-lain. Sedangkan bakteri yang menyebabkan hepatitis antara lain misal bakteri *Salmonella typhi*, *Salmonella paratyphi*, *tuberculosis*, *leptosvera*. Faktor non-infeksi misalnya karena obat. Obat tertentu dapat virus hepatitis C dapat menimbulkan problema pasca akut dan bahkan dapat terjadi kerusakan pada Hati. Untuk mengatasi masalah ini perlu adanya metode untuk mempermudah diagnosis yang mampu mengevaluasi keseluruhan hambatan dan kesulitan yang dialami pasien, maka dibutuhkan sebuah sistem yang mampu mendiagnosis penyakit hepatitis C menggunakan penerapan teknik data mining metode klasifikasi.

Data mining merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan dan machine learning untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial dan berguna yang tersimpan dalam database besar (Turban, dkk. 2005).

Berdasarkan banyaknya potensi penyebab munculnya Hepatitis yang berkaitan tentang bakteri, virus dan infeksi obat. Jawaban dari setiap pertanyaan gejala dan potensi terburuk kemudian dijadikan sebagai masukan untuk mengklasifikasi Hepatitis. Penerapan implementasi algoritma *MKNN* untuk

klasifikasi penyakit Hepatitis bertujuan untuk memudahkan ahli medis dalam memperoleh diagnosis sementara dengan mengklasifikasi gejala-gejala yang dialami oleh pasien dengan K data tetangga terdekat untuk mendapatkan nilai akurasi yang tinggi. *MKNN* merupakan algoritma yang dikembangkan dari algoritma *KNN*, algoritma *MKNN* menambahkan proses baru untuk melakukan klasifikasi yaitu perhitungan nilai validitas untuk mempertimbangkan validitas antar data latih dan perhitungan *weighted voting* untuk menghitung bobot dari masing-masing terdekat. Penambahan 2 proses baru dalam *MKNN* diharapkan dapat memperbaiki setiap kesalahan pada proses *KNN*.

Sebagai pendukung proses olah data, peneliti memutuskan menggunakan metode *Modified K-Nearest Neighbor (MKNN)*. Karena pada penelitian sebelumnya yang dilakukan oleh (Rissa Helilintar, dkk, 2017) dengan judul Perancangan sistem diagnosis penyakit hepatitis menggunakan metode *KNN*. Dari hasil penelitian dapat disimpulkan bahwa hasil akurasi terbaik adalah 78%. Dan pada penelitian yang dilakukan oleh (Imanda, dkk, 2018) dengan judul Klasifikasi Kelompok Varietas Unggul Padi Menggunakan *Modified K-Nearest Neighbor*. Pada penelitian ini varietasdirekayasa menggunakan *data mining* dan didapatkan nilai rata – rata akurasi maksimum sebesar 79,96%. Dengan melihat pada peneliti – penelitian sebelumnya diharapkan pada peneliti mampu menghasilkan akurasi yang lebih baik. Berdasarkan paparan diatas, penulis tertarik untuk mengimplementasikan deteksi dini hepatitis melalui penelitian dengan judul **“Diagnosis penyakit Hepatitis C berdasarkan faktor – faktor yang mempengaruhi menggunakan metode *MKNN (Modified K – Nearest Neighbor)*”** berdasarkan munculnya gejala – gejala Hepatitis yang berkaitan. Jawaban dari setiap pertanyaan gejala kemudian dijadikan sebagai masukan untuk mengklasifikasi penyakit Hepatitis. Penggunaan Metode Modified K – Nearest Neighbor diharapkan dapat menghasilkan nilai akurasi yang akurat sebagai tanda bahwa sistem layak untuk digunakan.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, rumusan permasalahan yang akan diselesaikan dalam penelitian tugas akhir ini yaitu:

- 1 Berapa k Optimal pada metode *Modified K-Nearest Neighbor* untuk diagnosis penyakit Hepatitis C?
- 2 Berapa tingkat Akurasi, Presisi dan Recall pada Metode K-Nearest Neighbor untuk mendeteksi penyakit Hepatitis C?

1.3. Batasan Masalah

Batasan Masalah yang digunakan meliputi sebagai berikut:

1. Data yang digunakan adalah data set bersumber dari “UCI – Machine Learning Repository Hepatitis C”, pada tahun 2019.
2. Data yang dijadikan kriteria meliputi 20 atributte yaitu, Age, Gender, BMI, Nausea, Headchea, Diarrhea, Fatigue & Generalize, Jaundice, Epigastric, WBC, RBC, HGB, Plat, AST1, ALT1, ALT4, ALT36, ALT48, RNA Base, RNA EF, Baseline histological.
3. Data yang digunakan yang mempunyai label kelas 1 dan 2 dalam data Hepatitis C.

1.4. Tujuan

Berdasarkan permasalahan yang telah dikemukakan sebelumnya, maka dapat ditarik tujuan dari penelitian yaitu:

1. Untuk mengetahui berapa K-Optimal pada metode Modified K-Nearest Neighbor untuk diagnosis penyakit Hepatitis C.
2. Untuk mengetahui berapa tingkat Akurasi, Presisi, Recall pada Metode K-Nearest Neighbor untuk mendeteksi penyakit Hepatitis C.

1.5. Manfaat Penelitian

Manfaat dari penelitian ini adalah:

1. Bagi umum manfaat yang diperoleh dari peneliti ini yaitu mampu memberi informasi mengenai indikator apa saja yang dapat antara pasien yang terindenfikasi penyakit Hepatitis dan yang tidak terkena penyakit Hepatitis.
2. Bagi peneliti dapat mengetahui dan menerapkan data mining metode K – Nearst Neighbor yang mempunyai akurasi tinggi untuk mengklasifikasi penyakit Hepatitis C yang sudah dipelajari.
3. Bagi peneliti lain diharapkan dapat menjadi metode acuan yang dapat dikembangkan dengan optimal.

