

BAB I

PENDAHULUAN

1.1 Latar Belakang

Penyakit TBC paru masih merupakan permasalahan serius dalam skala global. Menurut laporan dari WHO tahun 2017, TBC paru menyebabkan sekitar 1,3 juta kematian, dengan tambahan 300.000 kematian yang terjadi pada pasien TBC paru yang juga terinfeksi HIV. Indonesia, sebagai negara dengan peringkat ketiga tertinggi setelah India dan Cina dalam jumlah kasus TBC paru, tercatat sebagai penyumbang signifikan dalam statistik ini. Dalam dua per tiga dari seluruh kasus TBC di seluruh dunia, delapan negara utama terlibat, di mana India menyumbang sebanyak 27%, Cina 9%, dan Indonesia 8%. Selain itu, prevalensi TBC paru di Indonesia terbagi menjadi tiga wilayah, yaitu Sumatera dengan 33%, Jawa dan Bali dengan 23%, dan Indonesia bagian timur dengan 44%[1]. Penyakit Tuberkulosis Paru (TB Paru) adalah penyakit disebabkan oleh infeksi bakteri *Mycobacterium tuberculosis*. TB Paru dapat menyebar melalui udara dan dapat menginfeksi sistem pernapasan manusia. Oleh karena itu, deteksi dini dan diagnosis yang akurat sangat penting untuk mengendalikan penyebaran penyakit ini dan memberikan perawatan yang tepat kepada pasien[2]. TBC paru masih menjadi masalah kesehatan didunia, Indonesia menduduki posisi ketiga jumlah penderita TBC paru setelah India dan Cina. Kasus TBC paru lebih banyak dinegara berkembang contohnya Indonesia. Pemanfaatan salah satu teknologi informasi yang berdampak besar dalam aktivitas manusia yang mana dapat memberikan banyak kemudahan dan solusi untuk menyelesaikan permasalahan yang ada dalam aktivitas manusia yaitu peran teknologi informasi *data mining* penerapan algoritma *K-Nearest Neighbor* untuk klasifikasi penyakit TBC paru, dimana algoritma *K-Nearest Neighbor* dimanfaatkan untuk mengklasifikasi dan meningkatkan diagnosis dini penyakit TBC paru berdasarkan gejala-gejala yang timbul dengan cepat dan tepat[1].

Metode yang dapat digunakan untuk membantu dalam diagnosis TBC Paru adalah dengan menggunakan metode klasifikasi. Klasifikasi adalah teknik yang digunakan untuk mengelompokkan data baru ke dalam kategori yang telah ditentukan sebelumnya. Salah satu algoritma klasifikasi yang digunakan adalah *K-Nearest Neighbor (K-NN)*. *K-Nearest Neighbor (K-NN)* adalah metode klasifikasi yang bekerja dengan mencari sekelompok k objek dari data pelatihan yang paling mendekati objek dari data baru atau data uji[3]. Algoritma *K-Nearest Neighbor* adalah algoritma yang sederhana, walaupun algoritma ini tergolong algoritma yang sederhana tetapi kinerja dari algoritma *K-Nearest Neighbor* sangat baik.

Kelebihan Algoritma *K-Nearest Neighbor* dilihat dari perbandingan metode dengan penelitian sebelumnya dapat dijabarkan seperti berikut. Penelitian sebelumnya telah secara luas menguji penggunaan *Algoritma K-Nearest Neighbor* dalam klasifikasi penyakit, dan metode ini terbukti efektif dalam mengklasifikasikan penyakit, Algoritma *K-Nearest Neighbor* dapat mengklasifikasikan balita stunting dengan mencari kelompok *testing*. Tujuan dari penelitian ini adalah untuk menilai tingkat akurasi yang paling tinggi dalam hasil klasifikasi menggunakan metode *K-Nearest Neighbor* terhadap data anak balita yang mengalami stunting. Hasil akurasi tertinggi dari penelitian menggunakan algoritma *Knearest Neighbor* adalah 95,00%[3]. Penelitian terkait selanjutnya yang berhubungan dengan klasifikasi penyakit menggunakan algoritma *Knearest Neighbor* adalah melakukan klasifikasi penyakit ginjal kronis menggunakan metode *K-Nearest Neighbor (KNN)* berdasarkan 24 gejala dari setiap 400 individu. Hasil akurasi dari penelitian yang dilakukan adalah 85,83%[4]. Penelitian terkait yang selanjutnya yang berhubungan dengan klasifikasi penyakit menggunakan algoritma *Knearest Neighbor* adalah Penerapan algoritma *K-Nearest Neighbor (K-NN)* dalam klasifikasi pasien penyakit jantung menghasilkan kinerja yang sangat baik. Hasil dari penelitian ini dengan penggunaan nilai $K = 3$ dan pembagian bobot data training sebesar 80%, 20% sisanya digunakan sebagai data testing.

Model *K-Nearest Neighbor* ini mencapai skor akurasi sebesar 92%, menunjukkan bahwa kemampuan model dalam mengklasifikasikan pasien penyakit jantung sangat tinggi. Presisi model mencapai 90% serta nilai recall sebesar 92% menandakan kemampuan model dalam mengidentifikasi pasien penyakit jantung dengan benar[5].

Penelitian yang saya ajukan berbeda dengan penelitian sebelumnya, perbedaannya yaitu penyakit yang di klasifikasi sebelumnya yaitu berupa stunting pada balita, penyakit ginjal kronis dan diabetes, sedangkan pada penelitian ini penyakit yang di klasifikasi yaitu TBC (*tuberculosis*). hal yang membedakan lainnya itu pada jumlah kelas yang akan diuji dimana penelitian ini terdapat 4 kelas sedangkan penelitian sebelumnya kebanyakan menggunakan 2 kelas dan 3 kelas. Data diambil langsung dari Puskesmas Bumiayu melalui wawancara dengan kepala Puskesmas Bumiayu Bapak dr. Ali Budiarto. Kasus TBC juga terjadi di wilayah Brebes, wilayah Brebes termasuk daerah dengan jumlah kasus yang tinggi terutama di wilayah Bumiayu. Kasus TBC paru di Bumiayu terbilang tinggi serta penularan penyakit TBC paru sangat mudah sehingga dapat dengan cepat menular kepada orang yang sehat. Penularan yang mudah dan cepat dapat menyebabkan bertambahnya penderita TBC paru lebih cepat dan memiliki resiko kematian yang tinggi sehingga perlunya diagnosis dini untuk mengurangi penularan dan mengurangi resiko kematian. Masalah fenomena tersebut dapat ditarik kesimpulan bahwa solusi untuk masalah tersebut telah banyak dilakukan pencarian solusinya tidak hanya dalam bidang kesehatan saja akan tetapi dalam bidang teknologi pun turut ikut andil, yaitu peran teknologi informasi *data mining* penerapan algoritma *K-Nearest Neighbor* untuk klasifikasi penyakit TBC paru dengan studi kasus di Puskesmas Bumiayu.

Puskesmas Bumiayu merupakan puskesmas yang terletak di wilayah Kabupaten Brebes dimana letaknya berada di Kecamatan Bumiayu, merupakan salah satu puskesmas dengan jumlah kasus TBC paru yang tinggi sehingga perlu dilakukan peningkatan diagnosis yang cepat dan akurat di Puskesmas Bumiayu adalah cara untuk pengendalian kasus TBC paru.

Kesimpulan dari uraian penjelasan yang sudah dijabarkan diatas, dapat disimpulkan bahwa Algoritma *K-Nearest Neighbor* sangat cocok digunakan untuk klasifikasi penyakit berdasarkan gejala yang terdapat pada dataset. Hasil akurasi ketika menggunakan Algoritma *K-Nearest Neighbor* dilihat dari penelitian sebelumnya mencapai 95%, dengan demikian dari uraian penjelasan diatas peneliti terdorong untuk mengambil skripsi yang berjudul Penerapan algoritma *K-Nearest Neighbor* untuk klasifikasi penyakit TBC paru di Puskesmas Bumiayu.

1.2 Rumusan Masalah

Berdasarkan latar belakang permasalahan diatas, maka diperoleh rumusan masalah pada penelitian ini yaitu mengukur tingkat akurasi dari hasil klasifikasi metode *K-Nearest Neighbor* dengan menggunakan *Confusion Matrix* pada data pasien TBC paru di Puskesmas Bumiayu?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengetahui tingkat akurasi dari hasil klasifikasi metode *K- Nearest Neighbor* terhadap data pasien TBC paru di Puskesmas Bumiayu.

1.4 Manfaat Penelitian

Adapun manfaat penelitian pada penelitian ini adalah sebagai berikut:

1. Penelitian ini diharapkan dapat memberikan manfaat bagi teori penelitian teknologi informasi dibidang data mining terutama yang berkaitan dengan klasifikasi pasien TBC paru di Puskesmas Bumiayu menggunakan algoritma *K- Nearest Neighbor*.
2. Penelitian ini diharapkan dapat terus berkembang untuk membantu pengembangan dalam penerapan algoritma *K- Nearest Neighbor* untuk klasifikasi pasien TBC paru di puskesmas Bumiayu.

1.5 Batasan Masalah

Adapun batasan-batasan masalah pada penelitian ini adalah sebagai berikut:

1. Data penelitian yang digunakan yaitu pasien TBC paru di puskesmas Bumiayu.
2. Algoritma yang digunakan untuk klasifikasi adalah algoritma *K-Nearest Neighbor*
3. Perhitungan akurasi menggunakan *Confusion Matrix*
4. Hasil dari penelitian ini berupa akurasi hasil klasifikasi metode *K-Nearest Neighbor* menggunakan *Confusion Matrix* terhadap data pasien TBC paru di Puskesmas Bumiayu.

1.6 Sistematika Penulisan

Adapun sistematika penulisan proposal ini adalah sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini merupakan bagian yang membahas latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Dalam bab ini, terdapat pembahasan mengenai berbagai teori yang mendukung proposal skripsi ini.

BAB III METODE PENELITIAN

Bab ini berisi terkait metode penelitian yang diterapkan, mencakup tahapan pengumpulan data, kerangka teori, dan prosedur dari metode yang diusulkan.

BAB IV HASIL DAN PEMBAHASAN

Dalam bab ini, penulis mengulas tentang proses pembuatan dan implementasi algoritma *K-Nearest Neighbor*, serta membahas hasil yang diperoleh dari penerapan yang dilakukan oleh penulis.

BAB V PENUTUP

Bab ini terdapat kesimpulan yang didapat dari penelitian yang telah dilaksanakan dan saran dari penulisan.

DAFTAR PUSTAKA

Halaman ini terdapat daftar referensi yang menjadi acuan dalam penyusunan proposal, meliputi bagian pendahuluan hingga metode penelitian.