

## ABSTRAK

Dalam dunia kesehatan mendiagnosis suatu penyakit adalah tindakan yang harus dilakukan sedini mungkin agar penyakit yang ditemukan pada tubuh pasien dapat segera diobati, sehingga tidak menimbulkan kematian. Penyakit Tuberkulosis (TB) merupakan penyakit yang dapat menimbulkan kematian jika tidak diketahui dan tidak diobati secara rutin. Oleh karena itu, untuk mengurangi angka kematian dari penderita Tuberkulosis, pakar kesehatan harus mendiagnosis penyakit TB sedini mungkin.

Dari data gejala utama, hasil pemeriksaan laboratorium dan hasil rontgen dapat dimanfaatkan untuk diolah menggunakan teknik penambangan data dengan menggunakan metode *Naïve Bayesian*. Metode *Naïve Bayesian* akan menghitung probabilitas untuk setiap nilai kejadian dari atribut target pada setiap kasus (sampel data). Selanjutnya, *Naïve Bayesian* akan mengelompokkan sampel data tersebut ke kelas yang mempunyai nilai probabilitas paling tinggi.

Keluaran dari sistem adalah sebuah identifikasi atau sebuah prediksi jenis TB yang diderita oleh pasien berdasarkan lokasi anatomi yang terserang TB yaitu TB Paru, TB Ekstra Paru dan tidak TB. Peneliti melakukan pengujian dengan data sebanyak 237 data dan menggunakan fold bernilai 3, 5, 7 dan 9. Pengujian dilakukan sebanyak 24 kali dengan rata-rata akurasi tertinggi sebesar 85,95 %.

## ABSTRACT

In the world of health diagnosing a disease is an act that should be done as early as possible in order for the disease were found can be treated immediately, so it won't cause a death. Tuberculosis (TB) is a disease that can cause a death if unrecognized or treated regularly. Therefore, to reduce the death rate of tuberculosis sufferers, the health experts have to diagnose it as early as possible.

Based on the main indication data, laboratory test results and the results of rontgen can be used to be proceed with the data mining techniques using *Naïve Bayesian method*. It will calculate the probability for each case values of the target attribute in every case (sample data). Next, this method will classify the sample data to the class which has the highest probability value.

The output of the system is an identification or a prediction of a type of Tuberculosis that suffered by the patients. Researcher conducted a testing with the data of 237 and using fold in 3, 5, 7 and 9. The tests were done 24 times with the highest average 85,95%.

