

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Analisis *Naive Bayes* dipakai untuk memprediksi suatu kejadian, misalnya: untuk memprediksi rating film, kelahiran pada data ibu hamil dan usia kelahiran. Beberapa penelitian yang sudah memakai *Naive Bayes* untuk dibahas sebagai berikut ini. Algoritma *Naive Bayes* juga dipakai untuk mendeteksi penyakit autisme serta jenis autisme pada anak. Dalam hal [1], model *Naive Bayes* dapat dijadikan sebagai penggolong yang baik dalam menentukan apakah suatu anak mempunyai gangguan autisme atau tidak. Penelitian [1] yang menggunakan model *Naive Bayes* adalah diagnosis anak autisme dengan 33 gejala dan 3 diagnosis (autisme rendah, autisme sedang, autisme berat). Penelitian [1] menentukan hasil klasifikasi *Naive Bayes* dengan hasil yang memiliki nilai terbesar antara hasil perhitungan 3 diagnosis tersebut. Algoritma *Naive Bayes* juga dipakai untuk mendeteksi pengirim *tweet*. Penelitian [2] akan mencoba menggunakan metode klasifikasi *Naive Bayes* dengan memanfaatkan fitur frekuensi kemunculan kata, n-grams, formalisasi kata, dan penggunaan huruf kecil semua maupun biasa. Setiap *tweet* yang diolah menggunakan bahasa Indonesia. Metode Klasifikasi *Naive Bayes* dipakai untuk menilai apakah sebuah SMS tertentu dikirim oleh orang yang sama atau tidak. Semua data SMS yang ada ditokenisasi lalu dihitung frekuensi kemunculannya. Fitur yang digunakan adalah pemanfaatan unigram dan bigram dari token yang diambil berdasarkan jumlah kemunculan angka ke-n terbanyak. Token tersebut lalu dijadikan dasar untuk pembentukan *training set* dan *test set*. Lalu dihitung nilai probabilitas yang paling tinggi antara dua kelas yang memungkinkan berdasarkan rumus teorema *Naive Bayes*. Algoritma *Naive Bayes* juga dipakai untuk pengenalan iris mata. Pada penelitian [3] klasifikasi *Naive Bayes* dipakai untuk mengetahui pengaruh klasifikasi terhadap tingkat akurasi pengenalan citra normalisasi iris dan citra annular iris. Algoritma *Naive Bayes* dalam penelitian ini dipakai untuk melakukan analisis *data mining* terhadap sejumlah artikel JUTISI untuk menentukan sitasi ke artikel JUTISI.

## 1.2 Rumusan Masalah

Dari latar belakang tersebut, maka rumusan masalah yang di dapat adalah :

1. Apakah analisis *Naive Bayes* dapat mengklasifikasikan artikel JUTISI terkait sitasinya?
2. Atribut apa yang menentukan sitasi terhadap artikel ?

## 1.3 Tujuan Pembahasan

Melakukan analisis data *mining* terhadap sejumlah artikel untuk menentukan klasifikasi ke artikel JUTISI berdasarkan sitasi.

## 1.4 Ruang Lingkup

Yang akan digunakan untuk melakukan analisis data *mining* terhadap sejumlah artikel JUTISI untuk menentukan sitasi JUTISI dengan metode:

1. Klasifikasi *Naive Bayes*.
2. Aplikasi yang digunakan untuk menganalisis data JUTISI : Weka.

## 1.5 Sumber Data

Sumber data digunakan adalah data yang berisi tentang list mengenai makalah JUTISI. Sumber data yang digunakan adalah data sekunder yang diperoleh dari website JUTISI dan Google Scholar.

## 1.6 Sistematika Penyajian

- Bab 1 : Pendahuluan  
Bab ini berisi tentang Latar Belakang, Rumusan Masalah, Tujuan Pembahasan, Ruang Lingkup, Sumber Data dan Sistematika Penyajian.
- Bab 2 : Kajian Teori  
Bab ini berisi tentang pembahasan *Naive Bayes*, Pre-prosesing dan Weka.
- Bab 3 : Analisis dan Desain  
Bab ini berisi tentang Data Sumber, Atribut yang akan diproses dan Pre-prosesing untuk data JUTISI.
- Bab 4 : Implementasi

Bab ini berisi tentang hasil *naïve bayes* dengan Bidang Kajian 1, 2, 3 dan Bidang Kajian Gabungan.

- Bab 5 : Pengujian

Bab ini berisi tentang perbandingan Data Baru 1 hanya dengan *mode test Cross-validation*. Data TA 1 versi 8 dengan *mode test Cross-validation, Percentage split, User supplied test* dan *Use Training Set*.

- Bab 6 : Simpulan dan Saran

Bab ini berisi tentang Simpulan dan Saran.

