

## BAB II

### TINJAUAN PUSTAKA

#### 2.1. Penelitian Terdahulu

Aida Indriani, 2014, *Klasifikasi Data Forum dengan menggunakan Metode Naïve Bayes Classifier*. Dalam penelitian ini. Klasifikasi adalah suatu proses pengkategorian yang dilakukan terhadap sekumpulan dokumen. Klasifikasi sangat penting untuk kemudahan pengguna dalam melakukan pencarian dokumen. Proses klasifikasi diawali dengan membagi koleksi dokumen menjadi 2 bagian utama yaitu data latih dan data uji. Dari data latih, dengan menggunakan metode tertentu diperoleh model klasifikasi yang nantinya akan digunakan untuk penentuan kelas terhadap data uji. Banyak sekali metode-metode yang dapat digunakan dalam klasifikasi dokumen, antara lain: *K-NN*, *Rocchio*, *WAK-NN*, *Naïve Bayes Classifier (NBC)*, dan lain-lain.

Ainalia Indranandita, Budi Susanto, Antonius Rachmat C, 2015, *Sistem Klasifikasi Dan Pencarian Jurnal Dengan Menggunakan Metode Naive Bayes Dan Vector Space Model*. Dalam penelitian ini dirancang supaya Klasifikasi Naive Bayes menghasilkan prediksi baik jika vektor yang terbentuk mewakili setiap kategori. Sedangkan pencarian *Vector Space Uoabt* dengan pendekatan *Cosrme* menghasilkan *recall* sebesar 54.8% dan *precision* sebesar 60.7%. Oleh karena itu, dibangun sistem klasifikasi dan pencarian yang dapat membantu pengguna, karena dilengkapi pencarian detil dengan pengetahuan label kategori hasil klasifikasi dan fitur metadata.

Mujib Ridwan, Hadi Suyono, dan M. Sarosa, 2014, *Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier*. Input dari sistem ini adalah data induk mahasiswa dan data akademik mahasiswa. Sampel mahasiswa angkatan 2005-2009 yang sudah dinyatakan lulus akan digunakan sebagai data training dan testing. Sedangkan data mahasiswa angkatan 2010-2011 dan belum lulus akan digunakan sebagai data target. Data input akan diproses menggunakan teknik data mining algoritma *Naive*

*Bayes Classifier (NBC)* untuk membentuk tabel *probabilitas* sebagai dasar proses klasifikasi kelulusan mahasiswa. *Output* dari sistem ini berupa klasifikasi kinerja akademik mahasiswa yang diprediksi kelulusannya dan memberikan rekomendasi untuk proses kelulusan tepat waktu atau lulus dalam waktu yang paling tepat dengan nilai *optimal*.

Sri Kusumadewi, 2016, *Klasifikasi Status Gizi Menggunakan Naive Bayesian Classification*. Penelitian ini menggunakan metode Naive Bayesian Classification (NBC). NBC merupakan algoritma klasifikasi yang sangat efektif (mendapatkan hasil yang tepat) dan efisien (proses penalaran dilakukan memanfaatkan input yang ada dengan cara yang relatif cepat). Algoritma NBC bertujuan untuk melakukan klasifikasi data pada kelas tertentu. Model statistik merupakan salah satu model yang terpercaya sangat andal sebagai pendukung pengambilan keputusan. Konsep probabilitas merupakan salah satu bentuk model statistik. Salah satu metode yang menggunakan konsep probabilitas adalah *Naive Bayesian Classification (NBC)*. Pada metode ini, semua atribut akan memberikan kontribusinya dalam pengambilan keputusan, dengan bobot atribut yang sama penting dan setiap atribut saling bebas satu sama lain.

Vidya Rizqiyani<sup>1</sup>, Anggraini Mulwinda, Riana Defi Mahadji Putri, 2017, *Klasifikasi Judul Buku dengan Algoritma Naive Bayes dan Pencarian Buku pada Perpustakaan Jurusan Teknik Elektro*. Berdasarkan Undang-undang Republik Indonesia Nomor 43 Tahun 2007 tentang Perpustakaan menyebutkan bahwa penyelenggaraan pendidikan nasional memerlukan pengembangan layanan perpustakaan berbasis teknologi informasi dan komunikasi sehingga Universitas Negeri Semarang khususnya Jurusan Teknik Elektro perlu adanya pengembangan layanan tersebut. Ketiadaan pengembangan layanan tersebut di perpustakaan Jurusan Teknik Elektro mengakibatkan pihak pengelola dan mahasiswa pengunjung perpustakaan mengalami kesulitan dalam mengelola dan menemukan referensi buku yang dicari. Pada penelitian ini, metode yang digunakan yaitu *waterfall* dengan pengujian *performance measure*, uji *black box*, dan uji sistem oleh ahli. Adapun pengujian *performance measure* memperoleh nilai *precision* 94,56%, *recall* 88,20%, *f-measure* 90,46%, dan akurasi 97,78%. Sedangkan uji sistem

oleh ahli menghasilkan *persentase* rata-rata 87,3%. Dari hasil-hasil tersebut, disimpulkan bahwa sistem berfungsi dengan baik serta algoritma *Naive Bayes* dapat digunakan untuk pengklasifikasian buku. Untuk penelitian selanjutnya, diperlukan penataan buku koleksi perpustakaan sesuai kategori dan penelitian ini dapat dikembangkan dengan algoritma klasifikasi lainnya untuk melakukan perbandingan hasil *performance measure*.

M.Zainudin, Erna Zuny Astuti, 2016, *Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Kelayakan Calon Nasabah Asuransi di Pt. BNI Semarang*. Metode penelitian yang dilakukan dalam penelitian ini antara lain, observasi, wawancara serta studi pustaka untuk mendapatkan data-data terkait dengan nasabah asuransi di PT. BNI Semarang. Klasifikasi data nasabah berdasarkan jenis kelamin, usia, status, pekerjaan, penghasilan, cara pembayaran premi dan masa pembayaran premi. Metode analisis sistem yang digunakan adalah metode *waterfall* dengan langkah-langkah, (1) Perencanaan, (2) Analisis, (3) Perancangan, (4) Implementasi, (5) Pemeliharaan.

Amir Hamzah, 2015, *Klasifikasi Teks Dengan Naive Bayes Classifier (NBC) Untuk Pengelompokan Teks Berita Dan Abstract Akademis*. Dalam penelitian ini penulis menggunakan filter frekuensi dokumen sebagai kriteria apakah suatu kata dapat dimasukkan dalam *vocabulary*. Filter frekuensi dokumen mempertimbangkan frekuensi dokumen yang memuat kata tersebut. Asumsinya semakin banyak dokumen dalam suatu kategori memuat suatu kata, maka kata tersebut dianggap semakin baik untuk dijadikan sebagai *feature* pembeda kategori dokumen tersebut dengan kategori dokumen yang lain.

Alfa Saleh, 2015, *Implementasi Metode Klasifikasi Naive Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga*. Pada penelitian sebelumnya metode *Naive Bayes* juga digunakan dalam memprediksi penyakit Dermatologi yang diabaikan tapi bahkan dapat menyebabkan kematian di mana metode *Naive Bayes* digunakan untuk mengenal pola data untuk mengungkap kemungkinan penyakit *dermatologi*[3]. Metode *Naive Bayes* juga dinilai berpotensi baik dalam mengklasifikasi dokumen dibandingkan metode pengklasifikasian yang lain dalam hal akurasi dan efisiensi komputasi [4].

Wiwik Muslehatin<sup>1</sup>, Muhammad Ibnu, Mustakim, 2017, *Penerapan Naïve Bayes Classification untuk Klasifikasi Tingkat Kemungkinan Obesitas Mahasiswa Sistem Informasi UIN Suska Riau*. Berdasarkan penelitian yang telah dilakukan sebelumnya menggunakan metode algoritma *Naïve Bayes Classification* dinyatakan bahwa metode ini cocok digunakan dalam kasus ini. Oleh karena itu dalam mengelaskan status kemungkinan obesitas pada mahasiswa sistem informasi uin suska riau dapat dilakukan menggunakan metode *naïve bayes classification*. Selain itu penelitian yang dilakukan oleh [11] tentang klasifikasi status gizi dengan hasil penelitiannya sebesar 93,2%. Menurut penelitiannya metode *Naïve bayes Classification* memiliki kelebihan dengan kinerja yang baik dengan hasil pengujian sebesar 93,2%. Adapun menurut [5] kekurangan dari metode ini adalah pada saat melakukan proses training dan testing klasifikasi tidak optimal jika tidak disimpan pada memori (*database*), untuk itu perlu disimpan hasil klasifikasi pada memori (*database*) untuk mengoptimalkan klasifikasi.

Effrida Manalu, Fricles Ariwisanto Sianturi, Mamed Rofendy Manalu, 2017, *Penerapan Algoritma Naïve Bayes Untuk Memprediksi Jumlah Produksi Barang Berdasarkan Data Persediaan Dan Jumlah Pemesanan Pada Cv. Papadan Mama Pastries*. Dengan teori diatas maka penulis menggunakan Metode *Naive Bayes* dengan bertujuan untuk mengantisipasi kelebihan dalam pemesanan atau penerimaan roti. Keuntungan penggunaan *Naive Bayes* adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. *Naive Bayes* sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan.

Tabel 2.1 Perbandingan Tinjauan Pustaka

| No. | Nama Peneliti  | Judul  | Fokus dan Tujuan Penelitian   |
|-----|--|--|---|
| 1.  | Aida Indriani  | <i>Klasifikasi Data Forum dengan menggunakan Metode Naïve Bayes Classifier</i>   | Dalam penelitian ini. Klasifikasi adalah suatu proses pengkategorian yang dilakukan terhadap sekumpulan dokumen. Klasifikasi sangat penting untuk kemudahan pengguna dalam melakukan pencarian dokumen.   |
| 2.  | Ainalia Indranandita, Budi Susanto, dan Antonius Rachmat C         | <i>Sistem Klasifikasi Dan Pencarian Jurnal Dengan Menggunakan Metode Naive Bayes Dan Vector Space Model</i>            | Dalam penelitian ini dirancang supaya Klasifikasi <i>Naive Bayes</i> menghasilkan prediksi baik jika vektor yang terbentuk mewakili setiap kategori.  |
| 3.  | Mujib Ridwan, Hadi Suyono, dan M. Sarosa                           | <i>Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier</i>    | Input dari sistem ini adalah data induk mahasiswa dan data akademik mahasiswa. Sampel mahasiswa angkatan 2005-2009 yang sudah dinyatakan lulus akan digunakan sebagai data training dan testing.  |
| 4.  | Sri Kusumadewi   | <i>Klasifikasi Status Gizi Menggunakan Naive Bayesian Classification</i>   | Algoritma <i>NBC (Naïve Bayes Classification)</i> bertujuan untuk melakukan klasifikasi data pada klas tertentu. Model statistik merupakan salah satu model yang terpercaya sangat andal sebagai pendukung pengambilan keputusan.   |
| 5.  | Vidya Rizqiyani1, Anggraini Mulwinda, dan Riana Defi Mahadji Putri | <i>Klasifikasi Judul Buku dengan Algoritma Naive Bayes dan Pencarian Buku pada Perpustakaan Jurusan Teknik Elektro</i> | Dari hasil-hasil tersebut, disimpulkan bahwa sistem berfungsi dengan baik serta algoritma <i>Naive Bayes</i> dapat digunakan untuk pengklasifikasian buku. Untuk penelitian selanjutnya, diperlukan penataan buku koleksi perpustakaan sesuai kategori dan penelitian ini dapat |

|    |  |  |   |
|----|--|--|---|
|    |  |  | dikembangkan dengan algoritma klasifikasi lainnya untuk melakukan perbandingan hasil <i>performance measure</i> .   |
| 6. | M.Zainudin, dan Erna Zuny Astuti               | <i>PENERAPAN ALGORITMA NAIVE BAYES UNTUK MENGLASIFIKASIKAN KELAYAKAN CALON NASABAH ASURANSI DI PT.BNI SEMARANG</i> | Metode penelitian yang dilakukan dalam penelitian ini antara lain, observasi, wawancara serta studi pustaka untuk mendapatkan data-data terkait dengan nasabah asuransi di PT. BNI Semarang. Klasifikasi data nasabah berdasarkan jenis kelamin, usia, status, pekerjaan, penghasilan, cara pembayaran premi dan masa pembayaran premi. |
| 7. | Amir Hamzah                                    | <i>Klasifikasi Teks Dengan Naïve Bayes Classifier (NBC) Untuk Pengelompokan Teks Berita Dan ABSTRACT Akademis</i>  | Dalam penelitian ini penulis menggunakan filter frekuensi dokumen sebagai kriteria apakah suatu kata dapat dimasukkan dalam <i>vocabulary</i> . Filter frekuensi dokumen mempertimbangkan frekuensi dokumen yang memuat kata tersebut.  |
| 8. | Alfa Saleh                                     | <i>Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga</i>      | Pada penelitian sebelumnya metode <i>Naive Bayes</i> juga digunakan dalam memprediksi penyakit Dermatologi yang diabaikan tapi bahkan dapat menyebabkan kematian di mana metode <i>Naive Bayes</i> digunakan untuk mengenal pola data untuk mengungkap kemungkinan penyakit dermatologi.  |
| 9. | Wiwik Muslehatin1, Muhammad Ibnu, dan Mustakim | <i>Penerapan Naïve Bayes Classification untuk Klasifikasi Tingkat Kemungkinan Obesitas</i>                         | Selain itu penelitian yang dilakukan oleh [11] tentang klasifikasi status gizi dengan hasil penelitiannya sebesar 93,2%. Menurut penelitiannya metode <i>Naive bayes</i>  |

|     |   |  |   |
|-----|---|--|---|
|     |   | <i>Mahasiswa Sistem Informasi UIN Suska Riau</i>   | <i>Classification</i> memiliki kelebihan dengan kinerja yang baik dengan hasil pengujian sebesar 93,2%.   |
| 10. | Effrida Manalu, Fricles Ariwisanto Sianturi, dan Mamed Rofendy Manalu | <i>Penerapan Algoritma Naïve Bayes Untuk Mempredesi Jumlah Produksi Barang Berdasarkan Data Persediaan Dan Jumlah Pemesanan Pada CV. PAPADAN MAMA PASTRIES</i> | Keuntungan penggunaan <i>Naive Bayes</i> adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. <i>Naive Bayes</i> sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan. |

