

BAB II. LANDASAN TEORI

2.1 Penelitian Terdahulu

Penelitian terdahulu ini menjadi salah satu acuan penulis dalam melakukan penelitian sehingga penulis dapat memperkaya teori yang digunakan dalam mengkaji penelitian yang dilakukan. Penulis mengangkat beberapa penelitian sebagai referensi dalam memperkaya kajian pada penelitian penulis. Penelitian terdahulu berupa beberapa jurnal terkait dengan penelitian yang akan dilakukan ditampilkan pada Tabel 2.1.

Tabel 2.1 Penelitian Terdahulu

No	Penelitian Terdahulu	
1.	Penulis	Riri Nada Devita, Heru Wahyu Herwanto, Aji Prasetya Wibawa
	Judul	Perbandingan Kinerja Metode Naive Bayes dan K-Nearest Neighbor untuk Klasifikasi Artikel Berbahasa Indonesia
	Jurnal	Jurnal Teknologi Informasi dan Ilmu Komputer, 2018
	Permasalahan	Kecocokan isi artikel dengan sebuah tema jurnal menjadi faktor utama diterima tidaknya sebuah artikel. Tetapi masih banyak mahasiswa yang bingung untuk menentukan jurnal yang sesuai dengan artikel yang dimilikinya.
	Metode	<i>Naive Bayes</i> dan <i>K-Nearest Neighbor</i>
	Hasil	Hasil dari penelitian tersebut menunjukkan hasil akurasi sebesar 70% diperoleh menggunakan metode <i>Naive Bayes</i> sedangkan metode <i>K-Nearest Neighbor</i> menghasilkan akurasi sebesar 40%.
2.	Penulis	Dea Herwinda Kalokasari, Dr. Imam Marzuki Shofi, Anif Hanifa Setianingrum
	Judul	Implementasi Algoritma Multinomial Naive Bayes Classifier pada Sistem Klasifikasi Surat Keluar
	Jurnal	Jurnal Teknik Informatika, 2018

	Permasalahan	Banyaknya kumpulan data surat keluar di DISKOMINFO Kabupaten Tangerang yang sulit diolah secara manual untuk diberi penomoran.
	Metode	Multinomial <i>Naive Bayes Classifier</i>
	Hasil	Hasil dari penelitian menunjukkan tingkat <i>accuracy</i> , <i>precision</i> , <i>recall</i> , dan <i>F-measure</i> dalam penerapan metode <i>Naive Bayes</i> berturut-turut sebesar 89,58%, 79,17%, 78,72%, dan 77,05%.
3.	Penulis	Admaja Dwi Herlambang, Satrio Hadi Wijoyo
	Judul	Algoritma Naive Bayes untuk Klasifikasi Sumber Belajar Berbasis Teks pada Mata Pelajaran Produktif di SMK Rumpun Teknologi Informasi dan Komunikasi
	Jurnal	Jurnal Teknologi Informasi dan Ilmu Komputer, 2019
	Permasalahan	Sumber belajar SMK Rumpun TIK perlu dikelompokkan berdasarkan kriteria atau ciri esensial setiap mata pelajaran produktif. Jika dilakukan secara manual membutuhkan waktu yang lama dan menghasilkan hasil yang kurang maksimal.
	Metode	<i>Naive Bayes</i>
	Hasil	Hasil dari penelitian tersebut diperoleh sembilan kelompok mata pelajaran produktif dan pengujian menghasilkan tingkat akurasi tertinggi sebesar 81,48%, sedangkan tingkat akurasi terendah sebesar 79,63% untuk penerapan metode <i>Naive Bayes</i> .
4.	Penulis	Hafiz Irsyad, Ahmad Farisi, Muhammad Rizky Pribadi
	Judul	Klasifikasi Opini Masyarakat Terhadap Jasa ISP MyRepublic dengan <i>Naive Bayes</i>
	Jurnal	Jurnal Nasional Teknik Elektro dan Teknologi Informasi, 2019
	Permasalahan	MyRepublic menyediakan media untuk berkomunikasi dalam memberikan pelayanan berupa konsultasi, tanya jawab, dan keluhan seputar jasa yang diberikan oleh

		MyRepublic menggunakan media sosial salah satunya adalah Twitter. Sulitnya mengklasifikasi keluhan serta tanya jawab yang disampaikan oleh pelanggan di Twitter.
	Metode	<i>Naive Bayes</i>
	Hasil	Hasil dari penelitian tersebut menunjukkan untuk penerapan metode <i>Naive Bayes</i> dihasilkan nilai akurasi positif sebesar 0,976%, negatif sebesar 0,828%, dan netral sebesar 0,833%, sehingga dihasilkan rata-rata sebesar 0,879%.
5.	Penulis	Zulfany Erlisa Rasjid, Reina Setiawan
	Judul	<i>Performance Comparison and Optimization of Text Document Classification using k-NN and Naïve Bayes Classification Techniques</i>
	Jurnal	<i>Procedia Computer Science, 2017</i>
	Permasalahan	Informasi yang bersifat tidak terstruktur pada dokumen korpus yang sulit diolah untuk dijadikan sebagai ringkasan.
	Metode	<i>K-Nearest Neighbor, Naive Bayes.</i>
	Hasil	Hasil dari penelitian menunjukkan metode <i>K-Nearest Neighbor</i> memperoleh nilai rata-rata akurasi sebesar 55,17% dan algoritma <i>Naive Bayes</i> sebesar 39,01%.
6.	Penulis	Veny Amilia Fitri, Rachmadita Andreswari, Muhammad Azani Hasibuan
	Judul	<i>Sentiment analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm</i>
	Jurnal	<i>Procedia Computer Science, 2019</i>
	Permasalahan	Kampanye anti-LGBT yang banyak didiskusikan oleh masyarakat Indonesia di media sosial Twitter. Sulitnya mengetahui komentar positif, negatif, dan netral pada diskusi tersebut jika dilakukan secara manual.
	Metode	<i>Naive Bayes, Decision Tree, Random Forest.</i>
	Hasil	Hasil penelitian menunjukkan hasil akurasi sebesar 86,43%

		untuk algoritma <i>Naive Bayes</i> jika dibandingkan dengan algoritma <i>Decision Tree</i> dan <i>Random Forest</i> yang memperoleh akurasi sebesar 82,91%.
7.	Penulis	Sitti Nurul Jannah Fitriyyah, Novi Safriadi, Enda Esyudha Pratama
	Judul	Analisis Sentimen Calon Presiden Indonesia 2019 dari Media Sosial Twitter Menggunakan Metode Naive Bayes
	Jurnal	Jurnal Edukasi dan Penelitian Informatika, 2019
	Permasalahan	Dalam menentukan sentimen mengenai pasangan Calon Presiden Indonesia 2019 melalui Twitter jika dilakukan secara manual membutuhkan usaha dan waktu yang cukup banyak.
	Metode	<i>Naive Bayes</i>
	Hasil	Hasil dari penelitian menunjukkan tingkat akurasi sebesar 64,6% dan 58% untuk pengujian 3 kelas paslon 01 dan 02, serta tingkat akurasi sebesar 77,7% dan 88% untuk pengujian 2 kelas paslon 01 dan 02.
8.	Penulis	Billy Gunawan, Helen Sasty Pratiwim Enda Esyudha Pratama
	Judul	Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes
	Jurnal	Jurnal Edukasi dan Penelitian Informatika, 2018
	Permasalahan	Sulitnya melakukan analisis terhadap sentimen pada ulasan produk melalui Twitter jika dilakukan secara manual.
	Metode	<i>Naive Bayes</i>
	Hasil	Hasil dari pengujian menunjukkan pada pengujian 3 kelas (negatif, netral dan positif) hasil terbaik didapatkan pada 90% data latih dan 10% data uji dengan nilai akurasi 77.78%, recall 93.33% dan precision 77.78% dan pada pengujian 5 kelas hasil terbaik didapatkan pada 90% data latih dan 10% data uji dengan nilai akurasi 59.33 %, recall 58.33 % dan precision 59.33 %.

9.	Penulis	Zamah Sari, Mochammad Sarosa, Suhari
	Judul	“Si Tole” Chatterbot untuk Melatih Rasa Percaya Diri Menggunakan Naive Bayes Classification
	Jurnal	Jurnal Nasional Teknik Elektro dan Teknologi Informasi, 2018
	Permasalahan	Melatih tingkat kepercayaan diri dengan menerapkan permainan berupa mengekspresikan pendapat dalam bentuk teks/ <i>chat</i> . Untuk mengetahui nilai tingkat percaya diri melalui hasil dari permainan tersebut perlu dilakukan proses klasifikasi.
	Metode	<i>Naive Bayes Classification</i>
	Hasil	Implementasi aplikasi Si Tole <i>Chatterbot</i> dapat dilakukan dengan baik. Respon dan data yang dikirimkan melalui Android sesuai dengan masukan atau keluaran yang diinginkan.
10.	Penulis	Rahmat Robi Waliansyah, Citra Fitriyah
	Judul	Perbandingan Akurasi Klasifikasi Citra Kayu Jati Menggunakan Metode Naive Bayes dan k-Nearest Neighbor (k-NN)
	Jurnal	Jurnal Edukasi dan Penelitian Informatika, 2019
	Permasalahan	Sulitnya melakukan pengelompokkan jenis kayu jati berdasarkan tekstur, berat, warna, dan lain sebagainya secara manual.
	Metode	<i>Naive Bayes</i> dan <i>K-Nearest Neighbor</i>
	Hasil	Hasil dari penelitian menunjukkan tingkat akurasi sebesar 82,7% pada algoritma <i>Naive Bayes</i> dan 70% oleh algoritma <i>K-Nearest Neighbor</i> .
	Penulis	Anita Desiani, Sugandi Yahdin, Desty Rodiah
	Judul	Prediksi Tingkat Indeks Prestasi Kumulatif Akademik Mahasiswa dengan Menggunakan Teknik <i>Data Mining</i>
	Jurnal	Jurnal Teknologi Informasi dan Ilmu Komputer, 2020
	Permasalahan	Melakukan prediksi tingkat

11.		Indeks Prestasi Kumulatif (IPK) mahasiswa. Hasil prediksi tersebut diharapkan dapat dijadikan sebagai dasar untuk menentukan tingkat kelulusan dan kualitas institusi pendidikan.
	Metode	Algoritma C4.5, <i>Naive Bayes</i>
	Hasil	Hasil dari penelitian menunjukkan tingkat akurasi sebesar 75,18% diperoleh algoritma C4.5 dan tingkat akurasi sebesar 74,47% diperoleh algoritma <i>Naive Bayes</i> .
12.	Penulis	Tri Herdiawan Apandi, Castaka Agus Sugianto
	Judul	Algoritma Naive Bayes untuk Prediksi Kepuasan Pelayanan Perekaman e-KTP
	Jurnal	Jurnal Informatika, 2019
	Permasalahan	Sulitnya mengetahui tingkat kepuasan masyarakat terhadap pelayanan e-KTP jika dilakukan secara manual
	Metode	<i>Naive Bayes, Decision Tree</i>
	Hasil	Hasil dari penelitian menunjukkan tingkat akurasi sebesar 91,70% untuk algoritma <i>Naive Bayes</i> dan algoritma <i>Decision Tree</i> sebesar 65,90%.
13.	Penulis	Firman Tempola, Miftah Muhammad, Amal Khairan
	Judul	Perbandingan Klasifikasi Antara KNN dan Naive Bayes pada Penentuan Status Gunung Berapi dengan K-Fold Cross Validation
	Jurnal	Jurnal Teknologi Informasi dan Ilmu Komputer, 2018
	Permasalahan	Belum diketahuinya perbandingan kinerja kedua algoritma <i>K-Nearest Neighbor</i> dan <i>Naive Bayes Classifier</i> dalam melakukan penentuan status gunung berapi.
	Metode	<i>Naive Bayes, K-Fold Cross Validation</i>
	Hasil	Hasil penelitian menunjukkan rata-rata akurasi sebesar 63,68% untuk algoritma <i>K-Nearest Neighbor</i> dan 79,71% untuk algoritma <i>Naive Bayes</i> .
	Penulis	Christian Cahyaningtyas, Hindriyanto Dwi Purnomo, Budhi Kristianto

14.	Judul	<i>The Use of Naive Bayes for Broiler Digestive Tract Disease Detection</i>
	Jurnal	<i>Journal of Information Technology and Computer Engineering</i> , 2018
	Permasalahan	Sulitnya melakukan klasifikasi untuk deteksi penyakit pada saluran pencernaan ayam broiler jika dilakukan secara manual
	Metode	<i>Naive Bayes</i>
	Hasil	Hasil dari penelitian menunjukkan tingkat akurasi sebesar 93,3% untuk penerapan metode <i>Naive Bayes</i>
15.	Penulis	Triana Dewi Salma, Yusuf Sulisty Nugroho
	Judul	Sistem Rekomendasi Pemilihan Sekolah Menengah Tingkat Atas Menggunakan Metode Naive Bayes
	Jurnal	Jurnal Ilmu Komputer dan Informatika, 2016
	Permasalahan	Permasalahan yang dihadapi oleh siswa untuk menentukan calon sekolah terbaik dan sesuai dengan kriterianya.
	Metode	<i>Naive Bayes</i>
	Hasil	Aplikasi rekomendasi berbasis web menggunakan metode <i>Naive Bayes</i> yang telah dibuat dapat membantu siswa untuk memilih sekolah terbaik berdasarkan kriteria-kriteria yang telah diajukan.

Dari beberapa judul penelitian yang telah dipaparkan, terpilih penelitian yang dilakukan oleh Riri Nada Devita, Heru Wahyu Herwanto, Aji Prasetya Wibawa yang berjudul “Perbandingan Kinerja Metode Naive Bayes dan K-Nearest Neighbor untuk Klasifikasi Artikel Berbahasa Indonesia” dipilih sebagai referensi utama penulis dalam melakukan penelitian. Hal ini dikarenakan adanya persamaan metode yang digunakan oleh penulis, yaitu *Naive Bayes*, sebagai metode utama dalam proses klasifikasi bidang penelitian dosen Jurusan Teknologi Informasi, Politeknik Negeri Malang berdasarkan judul publikasi ilmiah setiap dosen.

2.2 Dasar Teori

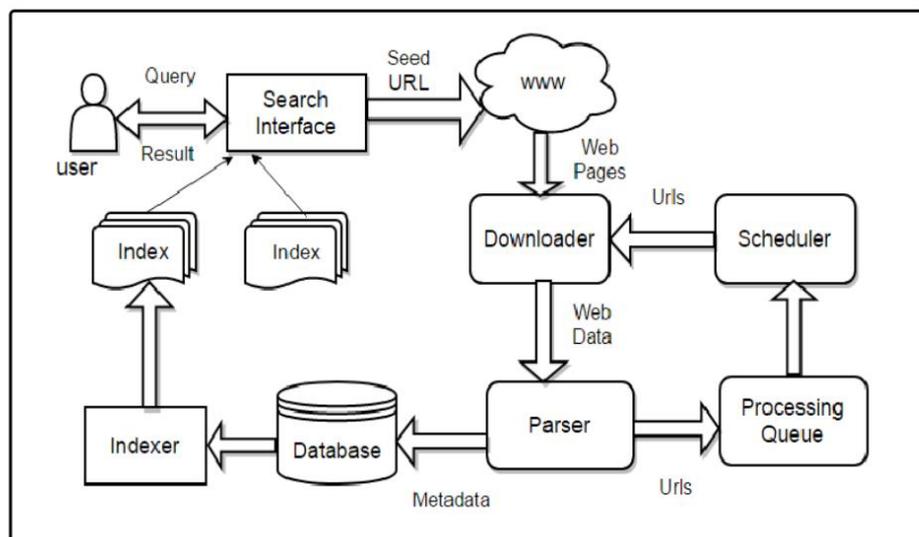
2.2.1 Sistem Informasi

Sistem informasi adalah suatu sistem di dalam organisasi yang mempertemukan kebutuhan pengolahan transaksi harian yang mendukung fungsi organisasi yang bersifat manajerial dengan kegiatan strategi dari suatu organisasi untuk dapat menyediakan laporan-laporan yang diperlukan oleh pihak luar tertentu (Oetari et al., 2012).

Pada penelitian ini sistem informasi yang dibuat akan menampilkan hasil klasifikasi judul publikasi dosen JTI POLINEMA serta kecenderungan bidang penelitian setiap dosen.

2.2.2 Web Crawling

Web crawling adalah penjelajahan rekursif yang dilakukan pada web untuk mengumpulkan setiap halaman yang terhubung dengan halaman inti. Proses *web crawling* dilakukan oleh program yang bernama *web crawler*. *Web crawler* mengambil halaman inti sebagai *input* dan mengembalikan halaman yang relevan dengan halaman inti sebagai *output* (Pavani & Sajeev, 2017).

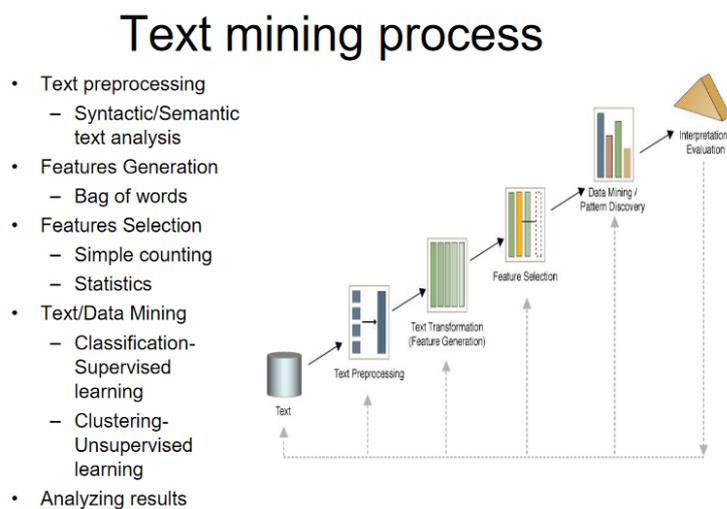


Gambar 2.1 Proses *Web Crawling*
(Sumber: (Pavani & Sajeev, 2017))

Pada penelitian ini diterapkan teknik *web crawling* untuk mengumpulkan dataset. Dataset tersebut berupa kumpulan judul publikasi berdasarkan bidang penelitian. Dataset diambil melalui website Google Scholar.

2.2.3 Text Mining

Text mining adalah proses penambangan data berupa teks yang didapatkan dari dokumen, tujuannya untuk mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisis keterhubungan antardokumen (Nindito, 2016). Target dari *text mining* berkisar dari literatur akademis hingga situs jejaring sosial, posting dan komentar mengenai berita, suara pelanggan, teks pidato, dan lain-lain (Jung & Lee, 2020). *Text mining* memungkinkan pengguna menganalisis dokumen teks tidak terstruktur diekstrak menjadi informasi yang bermakna (Choi et al., 2020). *Text mining* memiliki jumlah bidang yang tak terhitung jumlahnya, diantaranya adalah karakterisasi gagasan, pengenalan spam, dan *point displaying* (Kolluri & Razia, 2020).



Gambar 2.2 Proses Text Mining

(Sumber: (Segall et al., 2009))

Pada penelitian ini diterapkan teknik *text mining* untuk mengekstrak dokumen teks tidak terstruktur menjadi informasi yang bermakna. Informasi tersebut berupa hasil klasifikasi judul publikasi berdasarkan bidang penelitiannya.

Hasil dari klasifikasi tersebut digunakan sebagai dasar untuk mengetahui kecenderungan bidang penelitian pada setiap dosen berdasarkan publikasinya.

2.2.4 K-Fold Cross Validation

K-fold cross validation atau disebut sebagai estimasi rotasi adalah teknik validasi model untuk menilai bagaimana hasil statistik analisis akan menggeneralisasi kumpulan data independen. Teknik ini digunakan untuk melakukan prediksi model dan memperkirakan seberapa akurat sebuah prediktif saat dijalankan (Tempola et al., 2018).

Penerapan *k-fold cross validation* dilakukan dengan membagi dataset sejumlah k -subset. Pada percobaan pertama, subset S_1 berperan sebagai data pengujian dan subset lainnya diperlakukan sebagai data pelatihan, pada percobaan kedua subset S_1, S_3, \dots, S_k menjadi data pelatihan dan S_2 menjadi data pengujian, dan seterusnya. Proses tersebut dilakukan sejumlah k kali perulangan (Bramer, 2007).

Penelitian ini menggunakan *k-fold cross validation* dengan nilai $k = 5$. *Dataset* yang digunakan sejumlah 1900 judul publikasi yang diambil melalui *website* Google Scholar. *Dataset* tersebut dibagi menjadi dua data, yaitu *data training* sejumlah 380 dan *data testing* sejumlah 1520. Proses pembagian *dataset* dilakukan sejumlah 5 kali perulangan.

Dataset tersebut dikumpulkan melalui proses *crawling* dan diolah menggunakan metode *Naive Bayes* untuk menghasilkan data model klasifikasi *Naive Bayes*. Data model klasifikasi *Naive Bayes* yang dihasilkan digunakan sebagai dasar untuk melakukan klasifikasi publikasi dosen JTI POLINEMA.

2.2.5 Preprocessing

Tahap *preprocessing* bertujuan untuk menyiapkan dokumen teks yaitu berupa kumpulan judul publikasi menjadi data yang siap diolah pada proses selanjutnya (Devita et al., 2018). Tahapan *preprocessing* yang dilakukan yaitu:

- *Tokenizing*

Tokenizing adalah tahapan untuk melakukan pemenggalan dan pemfilteran kata pada dokumen. Pemenggalan kata dilakukan menggunakan spasi sebagai pemisah antar kata. Pemfilteran dilakukan dengan membuang karakter tertentu, seperti tanda baca.

- *Case folding*

Case folding adalah tahapan untuk mengubah semua huruf kapital pada dokumen menjadi huruf kecil. Karakter selain huruf A-Z dianggap sebagai *delimiter*.

- *Stopword Removal*

Stopword removal adalah tahapan untuk membuang kata-kata umum (*common word*) pada dokumen. Contoh *stop words* untuk Bahasa Inggris seperti “of”, “the”. Sedangkan untuk Bahasa Indonesia seperti “yang”, “di”, “ke”.

- *Stemming*

Merupakan suatu proses yang terdapat dalam sistem *Information Retrieval* yang melakukan transformasi kata yang terdapat pada suatu dokumen ke kata dasarnya (*root word*) (Fitriyyah et al., 2019). Tahapan proses *stemming* adalah sebagai berikut:

1. Menghapus *inflection suffix*, yaitu “-lah”, “-kah”, “-ku”, “-mu”, atau “-nya”
2. Cek *prefix* dan *suffix* yang tidak diperbolehkan yaitu (“be-” , “-i”) (“di-”, “-an”), (“ke-”, “-i”, “-kan”), (“me-”, “-an”), (“se-”, “-i”, “-kan”)
3. Menghapus *derrivation suffix*, yaitu: “-i”, “-an”, “-kan”
4. Menghapus *derrivation prefix*, yaitu: “di-”, “ke-”, “se-”, “te-”, “be-”, “me-”, atau “pe-”

2.2.6 Term Weighting (TF-IDF)

Term weighting adalah tahapan untuk menentukan bobot suatu kata pada dokumen menggunakan metode *Term Frequence-Inverse Document Frequency* (TF-IDF). Metode TF-IDF merupakan cara untuk mencari bobot suatu kata (*term*) pada sebuah dokumen (Robertson, 2004). Metode TF-IDF menggabungkan dua cara untuk melakukan perhitungan bobot, yaitu dengan menghitung frekuensi kemunculan kata disetiap dokumen tertentu (TF) dan melakukan perhitungan *invers* terhadap frekuensi dokumen yang mengandung kata tersebut (IDF) (Setianingrum et al., 2018). Rumus perhitungan TF dan IDF dapat dilihat pada persamaan 2.1 dan 2.2 (Devita et al., 2018).

$$TF(d,t) = f(d,t) \quad (2.1)$$

$$IDF(t) = 1 + \log\left(\frac{Nd}{df(t)}\right) \quad (2.2)$$

Dimana:

$f(d,t)$: frekuensi kemunculan *term t* pada dokumen *d*

Nd : jumlah seluruh dokumen

$df(t)$: jumlah dokumen yang terdapat *term t*

Untuk menemukan nilai TF-IDF menggunakan persamaan 2.3

$$W_{d,t} = TF(d,t).IDF(t) \quad (2.3)$$

Dimana:

$W_{d,t}$: Bobot *term* ke-*t* terhadap dokumen *d*

2.2.7 Naive Bayes Classification

Naive Bayes merupakan metode pengklasifikasian probabilistik sederhana. Metode ini melakukan perhitungan sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. *Naive Bayes* merupakan metode *supervised document classification* yang berarti membutuhkan *data training* (Devita et al., 2018). Algoritma *Naive Bayes* menggunakan teorema *Bayes* yang ada pada teori probabilitas yang menganggap semua atribut pada setiap kategori tidak memiliki ketergantungan satu sama lain (Herlambang & Wijoyo, 2019). Tahapan dari perhitungan *Naive Bayes* adalah sebagai berikut:

1. Menghitung probabilitas kategori dokumen dengan cara membagi jumlah seluruh dokumen pada suatu kategori dengan jumlah seluruh dokumen *training*. Sesuai dengan persamaan 2.4.

$$p(c_j) = \frac{n(doc_j)}{n(sampel)} \quad (2.4)$$

Dimana:

$p(c_j)$: probabilitas dokumen kategori

$n(doc_j)$: jumlah seluruh dokumen pada suatu kategori

$n(sampel)$: jumlah seluruh dokumen *training*

2. Menghitung probabilitas kata pada setiap kategori dengan cara membagi nilai frekuensi kemunculan kata setiap kategori ditambah satu dengan penjumlahan antara jumlah kosakata dengan jumlah seluruh kata dalam dokumen pada kategori tertentu. Sesuai dengan persamaan 2.5.

$$p(t_i|c_j) = \left(\frac{W_{i,j}+1}{\sum_{i=1}^n W_i + |kosakata|} \right) \quad (2.5)$$

Dimana:

$p(t_i|c_j)$: probabilitas kata setiap kategori

$W_{i,j}$: Nilai pembobotan dari kata yang muncul setiap kategori

$\sum_{i=1}^n W_i$: Jumlah total pembobotan seluruh kata setiap kategori

$|kosakata|$: jumlah kata unik pada seluruh kategori

3. Membandingkan dan mencari nilai probabilitas terbesar c_{MAP} . Nilai probabilitas terbesar tersebut digunakan untuk klasifikasi data uji pada dokumen publikasi kedalam setiap kategori yaitu bidang penelitian. Sesuai dengan persamaan 2.6.

$$c_{MAP} = \underset{c_{\text{res}}}{\operatorname{argmax}} p(c_j) \prod_i p(t_i|c_j) \quad (2.6)$$

Pada penelitian ini diterapkan metode *Naive Bayes Classification* untuk melakukan klasifikasi judul publikasi berdasarkan bidang penelitiannya. Proses klasifikasi menggunakan *dataset*. Atribut pada *dataset* yaitu judul publikasi pada setiap kategori tidak memiliki ketergantungan satu sama lain.

2.2.8 Confusion Matrix

Confusion matrix adalah metode yang digunakan untuk melakukan evaluasi terhadap hasil klasifikasi yang dilakukan. Metode ini menyajikan hasil klasifikasi menggunakan matriks yang ditampilkan pada Tabel 2.2 (Devita et al., 2018).

Tabel 2.2 *Confusion Matrix*

Klasifikasi Sebenarnya	Diklasifikasikan Menjadi	
	+	-
+	True Positive	False Positive
-	False Negative	True Negative

True Positive adalah jumlah *record* positif yang berhasil diklasifikasikan sebagai positif. *False Positive* merupakan *record* positif yang salah diklasifikasikan menjadi negatif. *False negative* merupakan *record* negatif yang salah diklasifikasikan sebagai positif. *True Negative* adalah *record* negatif yang berhasil diklasifikasikan sebagai *record* negatif.

Metode pengujian *confusion matrix* menghasilkan perhitungan dengan 4 output, yaitu:

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (2.7)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (2.8)$$

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (2.9)$$

2.2.9 ReactJS

ReactJs merupakan *framework* atau kerangka kerja yang bersifat *open source* yang menggunakan *library javascript* untuk membuat *user interface*. *ReactJS* digunakan untuk menangani pengembangan aplikasi *single-page* dan aplikasi *mobile* (Nursaid et al., 2020). *ReactJS* memiliki keunggulan dimana kerangka kerja ini memberikan kecepatan, *simplicity*, dan *scalability* (Khuat, 2018).

Pada penelitian ini sistem informasi hasil klasifikasi kecenderungan bidang penelitian dosen JTI POLINEMA dibuat menggunakan *Framework ReactJS*.