

BAB II

LANDASAN TEORI

2.1 Data Mining

Data mining merupakan proses penggalian informasi dan pola yang bermanfaat dari data yang sangat besar. Data mining mencakup pengumpulan data, ekstraksi data, analisis data, dan statistik data. Data mining juga dikenal sebagai *knowledge discovery*, *knowledge extraction*, *data/pattern analysis*, *information harvesting*, dan lain-lain (Muflikhah, Ratnawati and MP, 2018).

Data mining juga merupakan proses logis untuk menemukan informasi yang berguna. Setelah ditemukan informasi dan pola dapat digunakan untuk alat pendukung dalam pengambilan keputusan dalam mengembangkan bisnis. Alat data mining dapat memberikan jawaban untuk berbagai pertanyaan yang terkait dengan bisnis dan terlalu sulit untuk diselesaikan. Data mining juga dapat digunakan untuk meramalkan tren masa depan yang memungkinkan pebisnis membuat keputusan yang efektif, proaktif, dan dinamis.

Definisi lainnya untuk data mining adalah suatu proses menganalisis pola data yang tersembunyi menurut berbagai perspektif untuk kategorisasi menjadi informasi yang berguna, yang dikumpulkan di area umum, data warehouse untuk analisis yang efisien, algoritma data mining, memfasilitasi pengambilan keputusan bisnis, dan informasi lainnya. Data mining menggunakan analisis matematika dalam mendapat atau menemukan pola dan kecenderungan dari data. Pada umumnya, pola ini sukar ditemukan oleh eksplorasi data secara biasa/tradisional, hal ini disebabkan oleh terlalu rumitnya hubungan antardata atau juga dapat disebabkan oleh data yang begitu besar.

Data mining bertujuan untuk menemukan pola yang sebelumnya tidak diketahui. Jika pola-pola tersebut telah diperoleh maka dapat digunakan untuk menyelesaikan berbagai macam permasalahan. Data mining saat ini juga telah menjadi suatu teknologi baru yang kuat dengan potensi besar untuk membantu perusahaan fokus pada informasi paling penting dalam data yang telah mereka kumpulkan tentang perilaku pelanggan dan pelanggan potensial mereka. Melalui data mining,

perusahaan dapat menemukan informasi dalam data yang begitu besar melalui pengolahan yang tepat dan efektif dengan berbagai metode yang ada dalam data mining sehingga secara sederhana data mining dapat digambarkan sebagai suatu pola atau model atau kaidah atau pengetahuan yang dihasilkan dari data mining, seperti Gambar 2.1.



Gambar 2.1 Model atau pengetahuan Merupakan *Output* Data Mining

Untuk menghasilkan model ataupun pengetahuan maka sangat berpengaruh dari bagaimana data atau informasi sebelumnya atau bagaimana sejarah data atau informasi terdahulu sehingga dapat diprediksi atau diestimasi informasi di masa yang akan datang. Makin bagus informasi masa lampau atau data-data sebelumnya maka biasanya hasil yang diperoleh memiliki keakuratan yang sangat baik.

Sebagai contoh, jika seorang manager supermarket ingin mengetahui apakah seseorang pelanggan ketika membeli telur apakah dia juga akan membeli minyak? Prosesnya berdasarkan histori datanya dapat dilihat berikut ini (Arhami and Nasir, 2020).

Historical Data

<i>BasketId</i>	<i>Eggs</i>	<i>Oil</i>	<i>Milk</i>	...
1	yes	yes	no	...
2	no	yes	yes	...
3	no	no	yes	...
4	no	yes	yes	...
5	yes	yes	no	...
6	yes	no	yes	...
7	no	no	no	...
8	yes	yes	yes	...
...

Data Mining

Pola/Model: Eggs ---> Oil: Confodence = 75%, Support 37%

2.2 Algoritma *Naive bayes*

Bayesian Classification (Klasifikasi *Naive bayes*) merupakan salah satu metode klasifikasi dalam data *mining* yang menggunakan perhitungan probabilitas dan statistik. Metode ini dikenalkan oleh ilmuwan Inggris Thomas Bayes. Fungsinya yakni memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya. Persamaan dari *naive bayes* (Huda, Awangga and Fathonah, 2020) adalah sebagai berikut :

$$P(C_i | X) = \frac{P(X | C_i) P(C_i)}{P(X)} \dots\dots\dots (1)$$

Karena $P(X)$ konstan untuk semua kelas, maka dimaksimalkan rumus :

$$P(C_i | X) = P(X|C_i)P(C_i) \dots\dots\dots (2)$$

Keterangan :

- X = Kriteria suatu kasus berdasarkan masukan
- C_i = Kelas solusi pola ke-i, dimana i adalah jumlah label kelas
- $P(C_i|X)$ = Probabilitas label kelas C_i dengan kriteria masukan X
- $P(X|C_i)$ = Probabilitas kriteria masukan X dengan label kelas C_i
- $P(C_i)$ = Probabilitas label kelas C_i

Klasifikasi dengan *naive bayes* bekerja berdasarkan teori probabilitas yang memandang semua fitur dari data sebagai bukti dalam probabilitas. Hal ini memberikan karakteristik *naive bayes* sebagai berikut (Huda, Awangga and Fathonah, 2020):

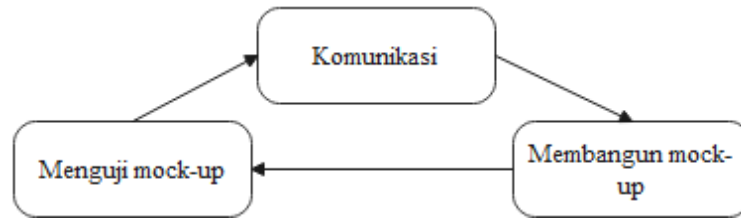
- a. Metode *naive bayes* bekerja teguh (*robust*) terhadap data-data yang terisolasi yang biasanya merupakan data dengan karakteristik berbeda (*outliner*). *Naive bayes* juga bisa menangani nilai atribut yang salah dengan mengabaikan data latih selama proses pembangunan model dan prediksi.
- b. Tangguh menghadapi atribut yang tidak relevan.
- c. Atribut yang mempunyai kolerasi bisa mendegradasi kinerja klasifikasi *naive bayes* karena asumsi independensi atribut tersebut sudah tidak ada.

Langkah-langkah algoritma naïve bayes pada pengadaan buku referensi perpustakaan sma negeri 1 trimurjo (Rahayu, S. and Sinaga, A.S. 2018):

- a. Persiapkan Pengkategorian Faktor Pendukung Pengambilan Keputusan
Sebelum memasuki tahap data *training*, maka perlu diketahui pengkategorian dari faktor-faktor yang mendukung pengadaan buku. Faktor yang mendukung dalam pengambilan keputusan tersiri dari tahun terbit, jumlah buku, dan peminatan.
- b. Persiapan Data *Training*
Tahapan ini dimulai dengan melakukan pengambilan data sampel pengadaan buku pustaka pada perpustakaan sekolah yang telah dikumpulkan yang akan dijadikan sebagai data *training*. Persiapan data yang akan dijadikan data *training* diambil dari data pengadaan buku perpustakaan sekolah sebelumnya (tidak disebutkan tahun). Data yang diambil sebanyak 12 data pengadaan buku pustaka. Berdasarkan hasil pengolahan data *training* tersebut, nantinya akan dikelola untuk membuat suatu hasil ketentuan yang akan digunakan oleh pustakawan perpustakaan sekolah untuk membantu menentukan buku pustaka/referensi berdasarkan hasil prediksi dalam pengambilan keputusan.
- d. Pengujian Sampel
Pengujian ini bertujuan untuk mengetahui kinerja dari algoritma *naïve bayes* dalam mengklasifikasi data kedalam kelas yang telah ditentukan. dihitung klasifikasi data dengan mengelola atribut atau masukan data yang sudah ditentukan sebelumnya yaitu tahun terbit, jumlah, dan minat menggunakan algoritma *naïve bayes*.

2.3 Metode Pengembangan Sistem

Penelitian ini menggunakan metode pengembangan sistem prototipe yang terdiri dari 3 (tiga) tahapan pengembangan sistem, diantaranya adalah tahap mendengarkan pelanggan, tahap membangun *mock-up*, dan tahap menguji *mock-up* (Rosa, A.S. and Shalahuddin, M. 2018). Kerangka penelitian dari metode pengembangan sistem prototipe yang dibangun adalah seperti pada Gambar 2.2.



Gambar 2.2 Tahapan Pengembangan Sistem Prototipe (Shy Destiny Deer, 2017)

Adapun langkah atau tahapan penggunaan metode prototipe dalam penelitian ini adalah (Rosa, A.S. and Shalahuddin, M. 2018) :

a. Komunikasi/Tahap Mendengarkan Pelanggan

Pada tahapan ini, peneliti berkomunikasi dengan pelanggan untuk mengetahui kebutuhan sistem yang akan dibuat. Peneliti mendengarkan keluhan pelanggan tentang pengelolaan data pustaka di perpustakaan SMA Negeri 1 Trimurjo. Hasil komunikasi berupa pengumpulan data terdiri dari observasi dan wawancara.

b. Tahap Membangun/Memperbaiki *Mock-Up*

Setelah pengumpulan data yang dibutuhkan didapatkan, maka selanjutnya merancang sistem pengadaan buku referensi perpustakaan. Perancangan sistem terdiri dari *use case diagram*, *activity diagram*, *database*, dan antarmuka sistem berbasis *web*. Setelah dilakukan perancangan, maka rancangan tersebut diimplementasikan ke dalam pemrograman *web*. Bahasa pemrograman yang digunakan adalah *Dart* dan penyimpanan data menggunakan *MySQL*.

c. Tahap Pelanggan Melihat/Menguji *Mock-Up*

Tahap ini dilakukan pengujian aplikasi dengan metode pengujian *black box* terhadap implementasi dari sistem pengadaan buku referensi perpustakaan yang dibuat. Pengujian yang dilakukan adalah menguji fungsional sistem apakah berjalan sesuai dengan semestinya.




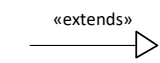
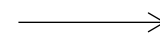
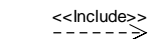
2.4 Pemodelan Sistem

2.4.1 Use Case Diagram

Use case mendeskripsikan sebuah interaksi antara satu atau lebih aktor dengan sistem informasi yang akan dibuat. *Use case* digunakan untuk mengetahui fungsi apa saja yang ada di dalam sebuah sistem informasi dan siapa saja yang berhak

menggunakan fungsi-fungsi itu. Penamaan pada *use case* simpel dan mudah dipahami (Rosa and Shalahuddin, 2018).


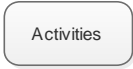




Tabel 2.1 Simbol *Use Case Diagram*

Keterangan	Simbol	Deskripsi
<i>Use Case</i>		Fungsionalitas yang disediakan sistem sebagai unit-unit yang saling bertukar pesan antar unit atau aktor; biasanya dinyatakan dengan menggunakan kata kerja diawal-awal frase nama <i>use case</i>
Aktor		Orang, proses atau sistem lain yang berinteraksi dengan sistem informasi yang akan dibuat di luar itu sendiri.
Asosiasi		Komunikasi antara aktor dan <i>use case</i> yang berpartisipasi pada <i>use case</i> atau <i>use case</i> memiliki interaksi dengan aktor.
Ekstensi		Relasi <i>use case</i> tambahan ke sebuah <i>use case</i> , dimana <i>use case</i> yang ditambahkan dapat berdiri sendiri walau tanpa <i>use case</i> tambahan.
Generalisasi		Hubungan generalisasi dan spesialisasi (umum-khusus) antara dua buah <i>use case</i> dimana fungsi yang satu adalah fungsi yang lebih umum dari lainnya.
<i>Include</i>		<i>Include</i> berarti <i>use case</i> yang ditambahkan akan selalu dipanggil saat <i>use case</i> tambahan dijalankan

2.4.2 Activity Diagram

Diagram aktivitas atau *activity diagram* menggambarkan *workflow* (aliran kerja) atau aktivitas dari sebuah sistem atau proses bisnis. Perlu diperhatikan disini adalah bahwa diagram aktivitas menggambarkan aktivitas sistem bukan apa yang dilakukan aktor, jadi aktivitas yang dapat dilakukan oleh sistem (Rosa and Shalahuddin, 2018).

Tabel 2.2 Simbol *Activity Diagram*

Simbol	Deskripsi
Status awal 	Status awal aktivitas sistem, sebuah diagram aktivitas memiliki sebuah status awal.
Aktivitas 	Aktivitas yang dilakukan sistem, aktivitas biasanya diawali dengan kata kerja.
Percabangan 	Asosiasi percabangan dimana jika ada pilihan aktivitas lebih dari satu.
Penggabungan 	Asosiasi penggabungan dimana lebih dari satu aktivitas digabungkan menjadi satu.
Swimlane 	Memisahkan organisasi bisnis yang bertanggung jawab terhadap aktivitas yang terjadi.
Status akhir 	Status akhir yang dilakukan sistem, sebuah diagram aktivitas memiliki sebuah status akhir.

2.5 Struktur Database

Database merupakan gabungan informasi yang memuat di dalam komputer secara terancang sehingga bisa dikontrol memakai suatu program komputer. Secara persepsi, *database* ialah gabungan dari data-data yang menata suatu arsip ataupun *file* yang saling berinteraksi dengan prosedur yang spesifik untuk memuat data baru maupun informasi (Subandi and Syahidi, 2018). Pengertian lain dari *database* yakni gabungan dari data yang saling berinteraksi antara satu dengan yang lainnya yang penataan menurut rancangan ataupun sistem tertentu (Rosa and Shalahuddin, 2018). Adapun struktur *database* terdiri dari :

a. Data

Sekumpulan fakta mengenai objek tertentu, orang dan lain-lain yang dinyatakan dengan angka, huruf, gambar, film, suara dan sebagainya yang relevan dan belum mempunyai arti.

b. Informasi

Hasil pengolahan data yang konkrit dan sudah mempunyai arti untuk mencapai suatu tujuan tertentu.

c. **Tabel**

Merupakan hal yang paling mendasar dalam hal penyimpanan data yang terdiri dari *field* dan *record*.

d. **Field**

Merupakan elemen dari tabel yang berisikan informasi tertentu yang spesifik tentang sub judul tabel pada sebuah item data. Syarat-syarat pembentukan Field Name pada tabel adalah:

1. Harus Unik atau Spesifik
2. Boleh disingkat

Pemisah sebagai pengganti spasi dalam pembentuk *field* adalah tanda lambang “_”. Contoh: Kode Barang menjadi KdBarang, KodeBrg, Kd_Brg, Kd_Barang.

Dalam sistem manajemen basisdata, terdapat tiga macam *field*:

1. Harus diisi (*required*)
2. Dapat diabaikan (*optional*)
3. Penghitungan dari field lainnya (*calculated*). Pengguna tidak dapat memasukkan data pada jenis *field* yang terakhir (*calculated*).

e. **Query**

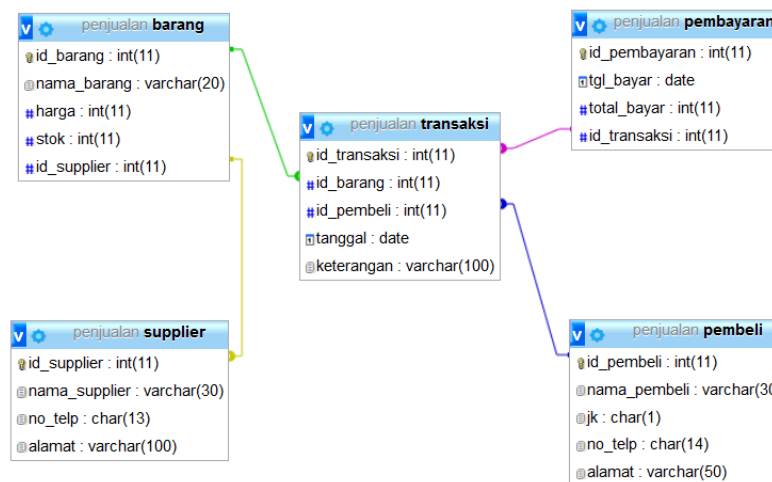
Query adalah pertanyaan atau permintaan informasi tertentu dari sebuah basis data yang ditulis dalam format tertentu. Terdapat tiga metode utama untuk membuat *query*, yaitu:

1. *Query by example* (QBE) adalah metode query yang disediakan sistem dalam bentuk *record* kosong dan pengguna dapat menentukan *field* dan nilai tertentu yang akan digunakan dalam *query*.
2. Bahasa *query* (*query language*) adalah bahasa khusus yang digunakan untuk melakukan *query* pada sebuah basisdata. Metode ini paling rumit tetapi paling fleksibel.

f. **Record**

Sekumpulan data yang saling berkaitan tentang sebuah subjek tertentu, misalnya data seorang siswa akan disimpan dalam *record* yang terdiri dari beberapa

kolom/*field*. Struktur *database* adalah cara data diorganisasi agar pemrosesan data menjadi lebih efisien. Sistem manajemen basis data (DBMS) adalah suatu aplikasi peranti lunak yang menyimpan struktur basis data-data itu sendiri, hubungan diantara data dalam basis data, dan nama-nama formulir, jenis-jenis data, angka dibelakang desimal, jumlah karakter, nilai-nilai *default* dan seluruh uraian *field* lainnya.



Gambar 2.3 Penerapan *Database* Penjualan Barang

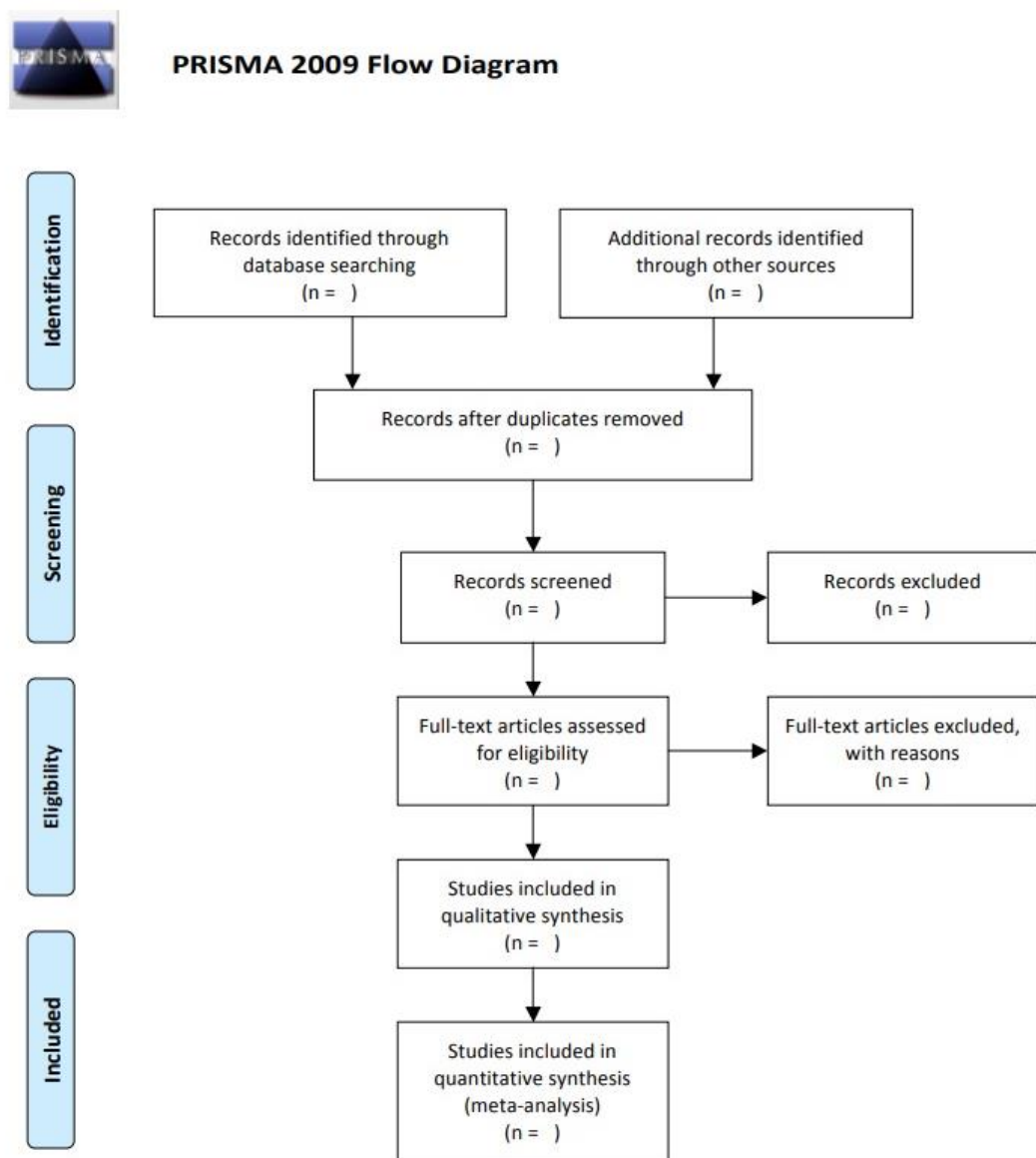
2.6 Black Box Testing

Black box testing merupakan pengujian untuk mengetahui fungsi perangkat lunak yang telah berjalan sesuai dengan kebutuhannya (Jurnal (ABDI) and Rahardja, 2020). Pengujian dimaksudkan untuk mengetahui apakah fungsi-fungsi, masukan, dan pengeluaran perangkat lunak sesuai dengan spesifikasi yang dibutuhkan (Rosa and Shalahuddin, 2018). *Black-box testing* akan berusaha dalam menemukan kesalahan dengan kategori sebagai berikut ini :

- Fungsi-fungsi yang rusak atau salah
- Kesalahan antarmuka perangkat lunak
- Kesalahan kinerja
- Kesalahan inisialisasi ataupun terminasi
- Kesalahan struktur data atau akses basis data

2.7 PRISMA Diagram

PRISMA (*Preferred Reporting Items for Systematic Reviews And MetaAnalyses*) merupakan alat dan panduan yang digunakan untuk melakukan penelitian terhadap sebuah *systematic reviews* dan *meta analysis*. Prisma diagram adalah diagram alir menggambarkan aliran informasi melalui fase yang berbeda dari tinjauan sistematis dengan memetakan jumlah catatan yang diidentifikasi, disertakan dan dikecualikan, dan alasan pengecualian (Page *et al.*, 2021).



Gambar 2.4 Diagram PRISMA

Melihat gambar di atas dapat di simpulkan bahwa pembuatan *Systematic Literature Review* (PRISMA) terdiri dari 4 langkah, yaitu: (1) identifikasi jurnal yang akan disertakan dalam meta-analisis (2) *screening*, penyaringan atau pemilihan data (3) kelayakan, menentukan artikel yang akan dijadikan bahan *literature review* dan (4) *included*, yakni penggabungan dan pelaporan hasil.

2.7.1 Identification

Identification (identifikasi) adalah kegiatan yang bertujuan untuk memeriksa dan menganalisa, dalam proses ini peneliti mengidentifikasi masalah penelitian (*research problem*) dikarenakan hal tersebut menentukan kualitas suatu penelitian. Dalam penelitian ini peneliti mengkaji permasalahan melalui banyak jurnal penelitian yang bersumber dari google cendikia. Kata kunci dari penelitian ini adalah penerapan algoritma *naive bayes*.

2.7.2 Screening

Screening (penyaringan) adalah penyaringan pemilihan data yang bertujuan untuk memilih masalah penelitian yang sesuai dengan topik yang diteliti. Adapun judul yang diteliti dalam penelitian ini adalah penerapan algoritma *naive bayes*. Dengan topik tersebut, data jurnal yang diakses dalam proses penelitian ini di-*screening* berdasarkan pada kriteria sebagai berikut :

- a. Jurnal diterbitkan dalam rentang waktu tahun 2018-2022
- b. Tipe jurnal *research articles*
- c. Jurnal yang dapat diakses secara penuh
- d. Pengecekan berdasarkan duplikasi (kesamaan jurnal)

2.7.3 Eligibility

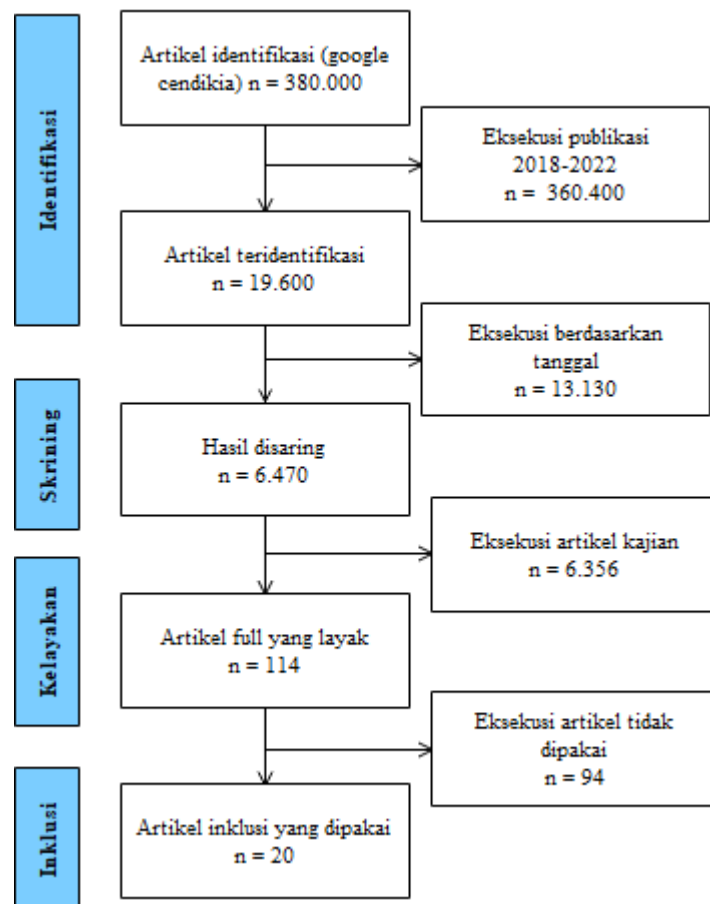
Tahap *eligibility* adalah tahap penentuan kelayakan artikel, peneliti menentukan artikel yang akan dijadikan bahan *literature review* yang telah menentukan kelayakan yaitu artikel harus berdasarkan riset asli dan artikel memiliki bahasan mengenai “*Naive bayes*”.

2.7.4 Included

Included merupakan tahap terakhir dari PRISMA (*Preferred Reporting Items for Systematic Reviews And Meta-Analyses*) *flow diagram*, pada tahap ini bertujuan untuk pengolahan data. Setelah artikel tersebut memenuhi kelayakan maka akan didapatkan jumlah artikel yang akan diteliti.

2.8 Tahapan Systematic Review

Ada beberapa tahapan *systematic review* yang harus dilakukan sehingga hasil dari studi literatur tersebut dapat diakui kredibilitasnya. Adapun tahapan-tahapan tersebut digambarkan pada Gambar 2.5.



Gambar 2.5 Diagram PRISMA: Tahap *Systematic Review*

Dalam melakukan *literature review* pada penelitian ini peneliti memasukkan kata kunci “*naive bayes*” dari google cendikia Indonesia dan melakukan beberapa tahapan yaitu identifikasi masalah yang dalam tahap ini peneliti mengumpulkan

artikel sebanyak 380.000 data yang berfokus pada algoritma *naive bayes* seperti Gambar 2.6.

The screenshot shows a Google Scholar search for "naive bayes". The search bar at the top contains "naive bayes" and a magnifying glass icon. Below the search bar, it says "Artikel" and "Sekitar 380 000 hasil (0,03 dtk)". On the left side, there are filters for "Kapan saja" (Since 2023, 2022, 2019), "Urutkan menurut relevansi" (Sort by relevance), "Urutkan menurut tanggal" (Sort by date), "Semua jenis" (All types), and "Artikel kajian" (Review articles). There are also checkboxes for "sertakan paten" (include patents), "mencakup kutipan" (include citations), and "Buat lansiran" (Create citations). The main results area shows several articles, including "Naïve Bayes" by G.I. Webb, E. Keogh, and R. Mikkilainen, "An empirical study of the naive Bayes classifier" by I. Rish, "The optimality of naive Bayes" by H. Zhang, and "A novel bayes model: Hidden naive bayes" by L. Jiang, H. Zhang, and Z. Cai.

Gambar 2. 6 Jumlah Hasil Pencarian *Naive Bayes*

Pada tahap ini juga dilakukan eksekusi publikasi berdasarkan tahun 2018-2022 dengan hasil teridentifikasi 19.600 data artikel seperti Gambar 2.7.

The screenshot shows the same Google Scholar search for "naive bayes", but with a date filter applied. The search bar still contains "naive bayes". Below the search bar, it says "Artikel" and "Sekitar 19.600 hasil (0,06 dtk)". The "Kapan saja" filter is now set to "2018" to "2022". The main results area shows articles filtered by year, including "Bayes' theorem and naive Bayes classifier" by D. Berrar, "An implementation of naive bayes classifier" by F.J. Yang, "Performance analysis of ANN and Naive Bayes classification algorithm for data classification" by M.M. Saritas and A. Yasar, and "A novel selective naive Bayes algorithm" by S. Chen, G.I. Webb, L. Liu, and X. Ma.

Gambar 2.7 Hasil Penyaringan Data Tahun 2018-2022

Screening yang dalam tahap ini peneliti mengumpulkan artikel sebanyak 6.470 data yang telah dilakukan tahap penyaringan dan memiliki fokus kajian algoritma *naive bayes* berdasarkan tanggal kajian seperti Gambar 2.8.

Google Cendekia "naive bayes"

Artikel Sekitar 6.470 hasil (0,20 dtk)

Kapan saja
Sejak 2023
Sejak 2022
Sejak 2019
Rentang khusus...

Urutkan menurut relevansi
Urutkan menurut tanggal

Semua jenis
Artikel kajian

sertakan paten
 mencakup kutipan

Buat lansiran

Artikel-artikel yang ditambahkan pada tahun lalu, diurutkan berdasarkan tanggal

PREDIKSI KASUS COVID 19 INDONESIA MENGGUNAKAN METODE NAÏVE BAYES
D Dafid, V Veratwi - JPP (Jurnal Kesehatan ...), 2022 - jurnal.poltekkespalembang.ac.id
2 hari yang lalu - ... **Naive Bayes**. Hasil: Prediksi yang dilakukan dengan menggunakan metode **Naive Bayes** ... Kesimpulan: Metode **Naive Bayes** mampu memberikan hasil prediksi ...
☆ Simpan Kutip

Analisis Sentimen Vaksinasi Booster Covid-19 pada Platform Twitter Menggunakan Metode Naive Bayes
DT Anggraeni - EXPERT: Jurnal Manajemen Sistem Informasi dan ..., 2022 - jurnal.ubl.ac.id
2 hari yang lalu - ... The method used is the **Naive Bayes** Classifier because the method is simple, the process is fast, and it has a fairly high level of confidence. In this method, public ...
☆ Simpan Kutip

Prediksi Predikat Kelulusan Mahasiswa Menggunakan Naive Bayes dan Decision Tree pada Universitas XYZ
A Wibowo, A Rohman - EXPERT: Jurnal Manajemen Sistem ..., 2022 - jurnal.ubl.ac.id
2 hari yang lalu - ... The method used in this study is the Cross Industry Standard Process for Data Mining CRISP-DM by utilizing the **Naive Bayes** algoritma in looking for graduate ...
☆ Simpan Kutip

Gambar 2. 8 Hasil Penyaringan Berdasarkan Tanggal Kajian

Setelah dilakukan tahap *screening* menghasilkan artikel kajian yang layak sebanyak 114 data seperti Gambar 2.9.

Google Cendekia "naive bayes"

Artikel Sekitar 114 hasil (0,04 dtk)

Kapan saja
Sejak 2023
Sejak 2022
Sejak 2019
Rentang khusus...

Urutkan menurut relevansi
Urutkan menurut tanggal

Semua jenis
Artikel kajian

Buat lansiran

Artikel-artikel yang ditambahkan pada tahun lalu, diurutkan berdasarkan tanggal

Sağlık Hizmetlerinde Güncel Makine Öğrenmesi Algoritmaları
LÖ POLATLI, MA KARADAYI - Eurasian Journal of Health ..., 2022 - dergipark.org.tr
3 hari yang lalu - UCI veri tabanından alınan kalp hastalığı verisine: **Naive Bayes** (Decision Tree (DT), K Nearest Neighbor (K-NN) ve Random Forest (RF) algoritmaları gibi denetimli ...
☆ Simpan Kutip

Machine Learning in Drug Metabolism Study.
K Sinha, J Ghosh, P C Sill - Current Drug Metabolism, 2022 - curoppmc.org
7 hari yang lalu - Metabolic reactions in the body transform the administered drug into metabolites. These metabolites exhibit diverse biological activities. Drug metabolism is the major ...
☆ Simpan Kutip 2 versi

Exploratory pharmacovigilance with machine learning in big patient data: a focused scoping review
BS Kaas-Hansen, S Gentile, A Caioli... - Basic & clinical ..., 2022 - Wiley Online Library
12 hari yang lalu - ... Two studies used artificial neural networks or **Naive Bayes** classifiers. One study used formal concept analysis for association mining, and another used temporal ...
☆ Simpan Kutip 4 versi

Gambar 2. 9 Hasil Penyaringan Kelayakan Artikel

Dari 114 data tersebut artikel yang yang tidak dipakai sebanyak 94 data dan artikel inklusi yang dipakai sebanyak 20 data sebagaimana terlihat pada Tabel 2.3.

Tabel 2.3 Penelitian Terkait

No	Judul, Penulis, Tahun	Jumlah & Atribut	Algoritma	Dataset	Validasi	Open Source Dataset	Akurasi
1	Jurnal Informatika Implementasi Algoritma <i>Naive bayes</i> Dalam Menentukan Konsentrasi Skripsi dan Rekomendasi Bahasa Pemrograman (Pratama and Yulmaini, 2018) https://jurnal.darmajaya.ac.id/index.php/JurnalInformatika/article/view/1031	14 variabel	<i>Naive bayes</i>	Data mahasiswa Teknik Informatika angkatan 2014 yang telah mengambil skripsi	-	Google Cendikia	96.5 %
2	Jurnal TEKNIKA Penerapan Data Mining dalam Perancangan Sistem Pendukung Keputusan Seleksi Penerimaan Beasiswa Menggunakan <i>Naive bayes</i> Classifier (Studi Kasus: IIB Darmajaya) (Agarina and Sutedi, 2020) http://repo.darmajaya.ac.id/3987/1/14_Penerapan%20Data%20Mining%20dalam%20Perancangan%20Sistem%20Pendukung%20Keputusan%20Seleksi%20Penerimaan%20Beasiswa%20Menggunakan%20Naive%20Bayes%20Classifier.pdf	NIS, Nama, Jenis Kelamin, Tempat Lahir, Tanggal Lahir, Jurusan yang dipilih,, Asal sekolah, Total Prestasi, Peringkat, Nilai Rata-rata dan Nilai Tes	<i>Naive bayes</i>	Data penerima beasiswa bidik misi tahun 2016-2020	-	Google Cendikia	-
3	Seminar Nasional Hasil Penelitian dan Pengabdian Sistem Pakar Diagnosa Penyakit Kulit Pada Balita Menggunakan Metode <i>Naive Bayes</i> dan <i>Forward Chaining</i> (Rahmatullah and Mawarni, 2020) https://jurnal.darmajaya.ac.id/index.php/PSND/article/view/2466	-	<i>Naive bayes</i> dan <i>Forward Chaining</i>	Data penyakit kulit pada balita	-	Google Cendikia	-
4	Skripsi Gabungan Repository IIB Darmajaya Implementasi Algoritma <i>Naive bayes</i> Dalam Klasifikasi Ban Terlaris Pada PT. Mitra Mekar Mandiri (Wijaya, 2018) http://repo.darmajaya.ac.id/1458/1/SKRIPSI%20FULL.pdf	-	<i>Naive bayes</i>	Data penjualan ban PT. Mitra Mekar Mandiri	-	Google Cendikia	Laris = 1151, Cukup Laris = 620 Kurang Laris = 629
5	Skripsi Gabungan Repository IIB Darmajaya	-	<i>Naive bayes</i> Classifier	Data organisasi mahasiswa	-	Google Cendikia	-

No	Judul, Penulis, Tahun	Jumlah & Atribut	Algoritma	Dataset	Validasi	Open Source Dataset	Akurasi
	Perancangan Sistem Pakar Pemilihan Organisasi Kemahasiswaan IIB Darmajaya Dengan Metode <i>Naive bayes</i> Classifier Berbasis Web (Pernando, 2019) http://repo.darmajaya.ac.id/1493/						
6	KRESNA : Jurnal Riset dan Pengabdian Masyarakat Implementasi Algoritma Naive Bayes Clasifier Untuk Mengelompokkan Naskah Berita Pendidikan dan berita Covid-19 (Mujahid, Syafrullah and Windarto, 2021) https://jurnalrpm.budiluhur.ac.id/index.php/Kresna/article/view/2	-	<i>Naive bayes</i> Classifier	Data berita khusus berbahasa Indonesia yang terdapat dalam Google Alerts dengan topik Kesehatan dan Pendidikan	-	Google Cendikia	74.58%
7	Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK) Algoritma Naive Bayes Untuk Klasifikasi Sumber Belajar Berbasis Teks Pada Mata Pelajaran Produktif Di SMK Rumpun Teknologi Informasi dan Komunikasi (Herlambang and Wijoyo, 2019) https://core.ac.uk/download/pdf/290149074.pdf	Animasi 2D, Animasi 3D Desain Multimedia, Basis Data, Pemodelan Perangkat Lunak, Pemrograman Dasar, Komputer Terapan Jaringan, Komunikasi Data, Sistem Operasi	<i>Naive bayes</i>	Data dokumen teks atau artikel teks yang diperoleh dari website yang berisikan materi-materi mata pelajaran produktif dari SMK Rumpun Teknologi Informasi dan Komunikasi	-	Google Cendikia	81,48%
8	Infotekjar : Jurnal Nasional Informatika dan Teknologi Jaringan Implementasi Algoritma <i>Naive bayes</i> Classifier Sebagai Sistem Rekomendasi Pembimbing Skripsi (Asfi and Fitrianiingsih, 2020) https://jurnal.uisu.ac.id/index.php/infotekjar/article/view/2536	Kompetensi, Jabatan Fungsional, dan Homebase Dosen	<i>Naive bayes</i> Classifier	Data history bimbingan skripsi beberapa tahun sebelumnya	-	Google Cendikia	pembimbing 1 sebesar 90% : 10% pembimbing 2 sebesar 30% : 70%

No	Judul, Penulis, Tahun	Jumlah & Atribut	Algoritma	Dataset	Validasi	Open Source Dataset	Akurasi
9	Prosiding Seminar Nasional Ilmu Komputer dan Teknologi Informasi Penerapan Metode Naïve Bayes pada Klasifikasi Judul Jurnal (Hardianti, Manga and Darwis, 2018) https://e-journals.unmul.ac.id/index.php/SAKTI/article/view/1838	-	Naïve Bayes	Data jurnal	-	Google Cendikia	50%
10	ILKOM Jurnal Ilmiah Klasifikasi Masyarakat Miskin Menggunakan Metode Naïve Bayes (Annur, 2018) https://jurnal.fikom.umi.ac.id/index.php/ILKOM/article/view/303	Umur, Pendidikan, Pekerjaan, Penghasilan, Tanggungan, Status (Kawin/Belum Kawin)	Naïve Bayes	Data masyarakat miskin dari tahun 2013-2015	-	Google Cendikia	73%
11	Jurnal Informatika Penerapan Algoritma Naïve Bayes untuk Rekomendasi Pakaian Wanita (Hayuningtyas, 2019) https://ejournal.bsi.ac.id/ejurnal/index.php/ji/article/view/4685	Style, Price, Rating, Size, Season, Neckline, Sleeve length, Waiseline, Material, Fabric type, Decoration, Patterntype, Recommendation	Naïve Bayes	Data pakaian wanita dari UCI machine learning repository	-	Google Cendikia	-
12	GEODIKA : Jurnal Kajian Ilmu dan Pendidikan Geografi Penerapan Naïve Bayes Dalam Mengklasifikasikan Masyarakat Miskin Di Desa Lepak, (Nurmayanti <i>et al.</i> , 2021) https://scholar.google.co.id/scholar?hl=id&as_sdt=0%2C5&q=Penerapan+Na%C3%AFve+Bayes+Dalam+Mengklasifikasikan+Masyarakat+Miskin+Di+Desa+Lepak&btnG=	-	Naïve Bayes	Data jumlah penduduk miskin tahun 2020	-	Google Cendikia	96,63%

No	Judul, Penulis, Tahun	Jumlah & Atribut	Algoritma	Dataset	Validasi	Open Source Dataset	Akurasi
13	Skripsi Gabungan Repository IIB Darmajaya Penerapan Metode Naïve Bayes Untuk Penentuan Siswa/i Terbaik Di SMA Negeri 1 Pardasuka (Wareka, 2019) http://repo.darmajaya.ac.id/1598/	UTS, UAS, Tugas, Absen, Sikap, dan Ekstrakurikuler	Naïve Bayes	Data siswa di SMA Negeri 1 Pardasuka tahun 2017/2018	-	Google Cendikia	-
14	Jurnal TIKOMSIN Penerapan Metode <i>Naive bayes</i> Untuk Klasifikasi Pelanggan (Putro, Vlandari and Saptomo, 2020) https://p3m.sinus.ac.id/jurnal/index.php/TIKomSiN/article/view/500	Jumlah Pembelian, Interval Waktu, dan Lokasi	Naïve Bayes	Data pembelian ayam	-	Google Cendikia	92%
15	JURSISTEKNI (Jurnal Sistem Informasi dan Teknologi Informasi) Penerapan Model Klasifikasi Metode <i>Naive bayes</i> Terhadap Penggunaan Akses Internet (Susana, Suarna and Kaslani, 2022) https://jursistekni.nusaputra.ac.id/article/view/96	Umur, Jenis Kelamin, Kelas, Tempat Tinggal, Gunakan HP, dan Gunakan Laptop	<i>Naive bayes</i>	Data keinginan konsumen	-	Google Cendikia	89.83%
16	Jurnal Infokar Sistem Pakar Deteksi Gizi Buruk Balita Dengan Metode Naïve Bayes Classifier (Sinaga and Simanjuntak, 2020) https://politeknikmeta.ac.id/meta/ojs/index.php/inkofar/article/view/110	Kwarshiorkor, Marasmik, dan Marasmus	Naïve Bayes Classifier	Data balita	-	-	-
17	Jurnal Teknologi Informasi & Komunikasi Digital Zone Penerapan Metode <i>Naive bayes</i> Dalam Pemilihan Kualitas Jenis Rumput Taman CV. Rumput Kita Landscape (Rahayu and Sinaga, 2018) http://journal.unilak.ac.id/index.php/dz/article/view/1942	Suhu Udara, Curah Hujan, Harga Pasaran, dan Kelembapan Udara	<i>Naive bayes</i>	Data klasifikasi budidaya jenis rumput	-	Google Cendikia	Rumput Golf= 0.4 Rumput Swiss= 0.4
18	KRESNA : Jurnal Riset dan Pengabdian Masyarakat Pemetaan Peningkatan Kinerja UMKM di Kabupaten Pesawaran Lampung Tengah Menggunakan Algoritma <i>Naive bayes</i> (Rohmawati and Lestari, 2021)	Nama Usaha, Jenis Usaha, Nama Produk, Izin yang Dimiliki, Aset,	<i>Naive bayes</i>	Data jenis usaha sembako dan kelontong sebanyak 305 UMKM di Kabupaten Pesawaran	-	Google Cendikia	98,68%

No	Judul, Penulis, Tahun	Jumlah & Atribut	Algoritma	Dataset	Validasi	Open Source Dataset	Akurasi
	https://scholar.google.co.id/scholar?hl=id&as_sdt=0%2C5&q=Pemetaan+Peningkatan+Kinerja+UMKM+di+Kabupaten+Pesawaran+Lampung+Tengah+Menggunakan+Algoritma+Naive+bayes&btnG=	Omset, dan Kriteria					
19	Jurnal Simetris Aplikasi Prediksi Usia Kelahiran Dengan Metode <i>Naive bayes</i> (Indraswari and Kurniawan, 2018) http://jurnal.umk.ac.id/index.php/simet/article/view/1827	Usia Ibu, Tekanan Darah, Jumlah bayi, Riwayat Persalinan, Riwayat Albortus Malnutrisi, Penyakit Lain, Masalah Penyakit Kehamilan Ini, Usia Kelahiran	<i>Naive bayes</i>	Data rekam medik pasien melahirkan di RSUD. Dr. Moewardi Provinsi Jawa Tengah dan Klinik Pratama An-Nisa selama tahun 2016	-	Google Cendikia	78,69%
20	Jurnal Gaussian Perbandingan Kinerja Metode Naïve Bayes dan K-Nearest Neighbor Untuk Klasifikasi Artikel Bahasa Indonesia (Devita, Herwanto and Wibawa, 2018) https://d1wqtxts1xzle7.cloudfront.net/72205112/pdf-libre.pdf?1633961390=&response-content-disposition=inline%3B+filename%3DPerbandingan_Kinerja_Metode_Naive_Bayes.pdf&Expires=1674567015&Signature=bIcfy2cS8GGNo573oC5dt7TIOXv7hUYMCx-x15VXkaj4uRN44Ww-lqKQ~GaJ~RhwFHe8acyVKcdWqx0d7t1~ogmCpJmkSuDIQZD-Rf3OWUdJMf1pvXZScXKC-Lb3nQU9~nFt4a7ma9FISgQ2nqaXODoqK3kyaSK~WPVupIE785Ofy8cZB1TkqXsgZZ8JnA1ieCBKqtISnSSNP2-8EulBlcutGK-LlWXwfnN5RSdmTcejC4EQdG-dDxppgryuFsBKGPZWDpZ~UDsrF9a7-NRk6WPq-wtWnGv4yOe7P0KpWcHidGV9wbZ~QT3r8R3ecl9OC60liq87FYLU4LJsDmHQ__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA	-	Naïve Bayes dan KNN	Data artikel bahasa Indonesia	-	Google Cendikia	Akurasi Naïve Bayes = 70% KNN = 40%

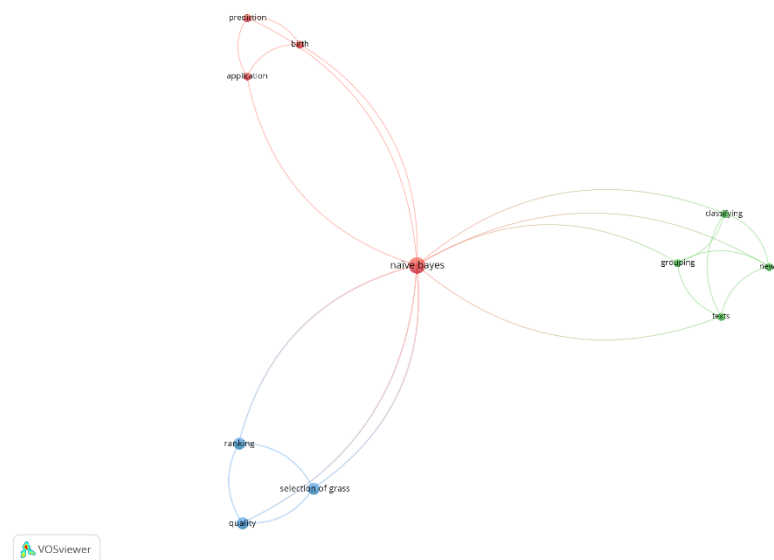
Adapun kesimpulan dari jurnal penelitian terkait adalah sebagai berikut :

- a. Algoritma *naïve bayes* dapat digunakan dalam memprediksi suatu pengambilan keputusan berdasarkan pelatihan dengan data yang sedikit, maka dari itu dapat digunakan dalam pengelolaan data secara *real time*.
- b. Algoritma *naïve bayes* mudah diimplementasikan pada aplikasi dengan media apapun seperti *Android* dan *web*.
- c. Algoritma *naïve bayes* dapat memprediksi multi kelas atau memprediksi probabilitas dari beberapa kelas variabel target.
- d. Algoritma *naïve bayes* bekerja sangat baik dengan variabel *input* kategori dibandingkan dengan variabel numerik.

2.9 VosViewer

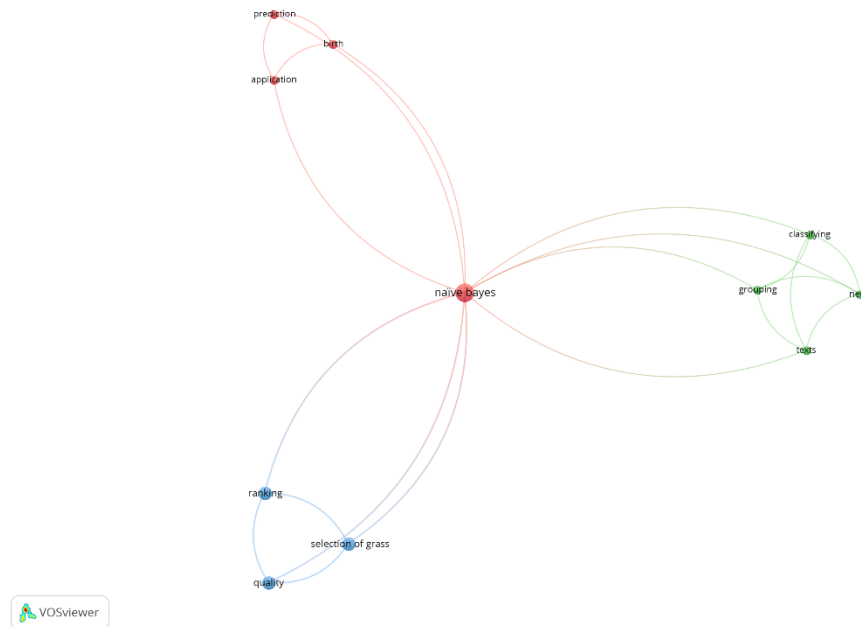
VOSviewer merupakan *software* yang dapat digunakan untuk membangun dan memvisualisasikan jaringan bibliometrik. Jaringan ini bisa berupa jurnal, peneliti, ataupun publikasi individu. Jaringan tersebut dapat dibangun berdasarkan kutipan, pasangan bibliografi, kutipan bersama, atau hubungan penulis bersama.

Berdasarkan hasil pencarian ke-15 judul artikel dengan subjek algoritma *sequential search*, selanjutnya dianalisis dengan *VOSviewer* dan didapat hasil sebagai berikut.



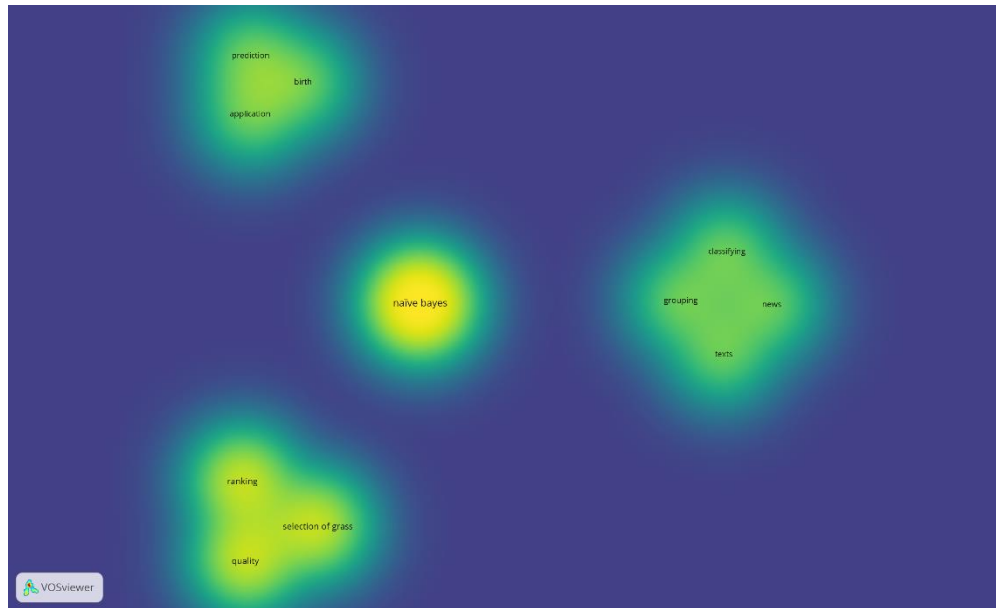
Gambar 2.10 Network Visualization

Gambar 2.10 di atas merupakan visualisasi dari hasil pengolahan data penelitian menggunakan kata kunci algoritma *sequential search*.



Gambar 2.11 *Overlay Visualization*

Gambar 2.11 adalah hasil analisis peneliti yang pernah melakukan penelitian terkait algoritma *naïve bayes*. Warna pada lingkaran di gambar menunjukkan bahwa semakin gelap warna lingkaran maka semakin tua usia penelitian dan semakin terang warna lingkaran maka semakin muda usia penelitian.



Gambar 2.12 *Density Visualization*

Gambar 2.12 di atas menunjukkan peta densitas yang merupakan hasil analisis yang menggunakan seluruh artikel algoritma *naive bayes*.

