

PENERAPAN METODE NAIVE BAYES UNTUK KLASIFIKASI KATEGORI OLAH PANGAN (STUDI KASUS DINAS KESEHATAN KOTA PALEMBANG)

Ajeng Oktaviyani*¹, Agustina Heryati², M.Fadhiel Alie*³

1,2,3 Fakultas Ilmu Komputer, Universitas Indo Global Mandiri
Email: *¹2020210018@students.uigm.ac.id, *²agustinaharyati@yahoo.com, ³fadhiel@uigm.ac.id
*Penulis Korespondensi

Abstrak

Dinas kesehatan merupakan bagian integral dari struktur pemerintahan yang memiliki tanggung jawab luas untuk mengelola berbagai aspek kesehatan, aktivitas pelayanan medis sehari-hari juga berperan dalam menjaga keseimbangan gizi yang sehat, yang tercermin dari makanan olahan yang baik dan berkualitas. Oleh karena itu perlu adanya pengelompokan dan pengklasifikasian pada olahan pangan. Adapun permasalahan yang terdapat dalam penelitian ini yaitu adanya kesulitan yang dialami oleh Dinas Kesehatan Kota Palembang dalam menentukan klasifikasi produk olah pangan. Oleh karena itu, peneliti menggunakan metode *Naive Bayes* dan *Software RapidMiner* sebagai perangkat lunak pendukung dalam penelitian ini. Dalam penelitian ini, pada pengujian *Accuracy* di dapatkan hasil nilai yang sama antara perhitungan *Rapidminer* dan perhitungan manual *Naive Bayes* yaitu kelas *positif* 94,52%: Nabati, pada pengujian *Precision* terdapat perbedaan antara perhitungan *RapidMiner* dengan perhitungan *Naive Bayes* manual, pada perhitungan *RapidMiner* didapatkan hasil yang lebih besar yaitu 93,22% Kelas *Positif*: Nabati, sedangkan perhitungan *Naive Bayes* manual didapatkan nilai 78%. Nilai serupa juga diperoleh pada pengujian *Recall* dimana antara perhitungan *Rapidminer* dan perhitungan manual *Naive Bayes*, nilainya sama yaitu 100% Kelas *positif*: Nabati, pada pengujian *F1 Score* diperoleh nilai yang sama antara perhitungan *Rapidminer* dan perhitungan manual *Naive Bayes* yaitu 100%, dan pada pengujian *Rapidminer Area Under The Curve (AUC)* Hasil yang diperoleh adalah 0,973 (kelas *positif*: Nabati), dan tingginya akurasi kurva ROC/AUC memberikan kriteria “*Excellent Clasifikasi* (klasifikasi yang sangat baik)”, sehingga penelitian ini nilai penggunaan bahan pangan nabati lebih mendominasi dibandingkan penggunaan bahan pangan hewani. Diharapkan hasil dari penelitian ini dapat membantu Dinas Kesehatan dalam mengklasifikasi kategori olah pangan.

Kata Kunci: *Naive Bayes*, Klasifikasi, Olah Pangan, Dinas Kesehatan Kota Palembang, Pengawasan Produk Pangan. Bahan Pangan, Hewani, Nabati.

Abstract

The Health Department is an integral part of the government structure with broad responsibilities for managing various aspects of health, including daily medical services activities that contribute to maintaining a healthy nutritional balance, reflected in good quality processed foods. Therefore, there is a need for grouping and classifying processed foods. The problem addressed in this research is the difficulty faced by the Palembang Health Department in determining the classification of processed food products. Hence, the researcher employed the Naive Bayes method and RapidMiner Software as supporting software tools in this study. The study found that in Accuracy testing, both RapidMiner and manual Naive Bayes calculations yielded the same result of 94.52% for the positive class: Plant-based. In Precision testing, there was a difference between RapidMiner and manual Naive Bayes calculations, with RapidMiner yielding a higher result of 93.22% for the positive class: Plant-based, while manual Naive Bayes calculation obtained a value of 78%. Similar values were also obtained in Recall testing, where both RapidMiner and manual Naive Bayes calculations yielded the same result of 100% for the positive class: Plant-based. In F1 Score testing, both RapidMiner and manual Naive Bayes calculations yielded the same result of 100%. In RapidMiner's Area Under The Curve (AUC) testing, the result obtained was 0.973 (positive class: Plant-based), and the high accuracy of the ROC/AUC curve indicates "Excellent Classification", suggesting that the use of plant-based food materials dominates over the use of animal-based food materials. It is hoped that the results of this research can assist the Health Department in classifying food processing categories.

Keywords : Naive Bayes, Classification, Processed Food, Palembang City Health Department, Food Product Supervision, Food Materials, Animal-based, Plant-based.

1. PENDAHULUAN

Produk olah pangan sangat berkaitan dengan kualitas sumber daya manusia. Produk olah pangan yang dikonsumsi menggambarkan status kesehatan gizi dan ekonomi seseorang. Oleh karena itu, perlu pemahaman tentang pembagian produk olah pangan secara umum agar bahan pangan tersebut dapat dimanfaatkan dan dikonsumsi sesuai dengan keperluan untuk mendapatkan zat gizi yang seimbang dan berkualitas.

Dinas Kesehatan juga merupakan bagian dari pemerintah yang bertanggung jawab atas berbagai aspek kesehatan masyarakat di suatu wilayah, Dalam kegiatan sehari-hari Dinas Kesehatan Kota Palembang bertanggung jawab dalam menjaga keseimbangan kesehatan gizi dan pangan yang ada di Kota Palembang serta memperhatikan produk olahan pangan apa saja yang berkualitas serta baik untuk di olah dan konsumsi oleh masyarakat.

Sedangkan menurut Devita, dkk (2018) Klasifikasi merupakan cara pengelompokan benda berdasarkan ciri – ciri yang dimiliki oleh objek klasifikasi. Dalam prosesnya, klasifikasi dapat dilakukan dengan banyak cara baik secara manual ataupun dengan bantuan teknologi. Dari empat belas variabel ini peneliti menggunakan tujuh variabel sebagai indikator dalam penelitian yaitu nama branding produk (CV/PT), jenis pangan, hasil pengolahan pangan, kemasan, ketahanan pangan, bahan pangan dan bentuk pangan menjadi indikator yang digunakan dalam riset produk olah pangan.

Berdasarkan latar belakang diatas maka penulis dapat merumuskan masalah yang ada yaitu Bagaimana Penerapan Metode Naive Bayes untuk Klasifikasi Kategori Olah Pangan (Studi Kasus Dinas Kesehatan Kota Palembang), serta Metode yang digunakan dalam penelitian ini yaitu *Naive Bayes* dengan olah data *Software RapidMiner* serta pembuatan Website menggunakan *Php, MySql, Visual Studio Code dan Xaampp*. Dengan data produk pangan Dinas Kesehatan Kota Palembang Tahun 2022 Data yang digunakan sebagai indikator terdiri dari nama branding produk (CV/PT), jenis pangan, hasil pengolahan pangan, kemasan, ketahanan pangan, bahan pangan dan bentuk pangan. Adapun tujuan dari penelitian ini adalah untuk mengklasifikasi atau mengelompokkan kategori produk olahan pangan hewani dan nabati di Dinas Kesehatan Kota Palembang guna memahami pola konsumsi masyarakat.

2. TINJAUAN LITERATUR

2.1. Olah Pangan

Menurut Fauzi (2019) pangan merupakan. bahan yang dimakan setiap hari oleh manusia yang berasal dari sumber air dan hayati yang merupakan kebutuhan dasar manusia yang paling utama. Serta menurut Maksu, dkk (2019) makna pangan adalah bahan yang dimakan hari demi hari dalam memenuhi kebutuhan untuk perkembangan, substitusi jaringan, pekerjaan, penunjang, dan pengaturan tindakan dalam tubuh.

Dari semua pendapat para ahli dapat disimpulkan bahwa olah pangan adalah komponen penting dalam kehidupan manusia yang mencakup berbagai produk dari berbagai sumber dan dapat diolah atau tidak. Ini memenuhi kebutuhan dasar kita, mendukung pertumbuhan dan fungsi tubuh, serta memberikan energi yang diperlukan untuk aktivitas sehari-hari.

2.2. Bahan Pangan Hewani

Bahan pangan hewani menurut Latifah (2020) Bahan pangan hewani adalah bahan pangan yang berasal dari hewan. Jenis bahan pangan hewani di antaranya adalah, ikan, daging telur, dan susu. Menurut Adi dan Hartati (2020) Bahan pangan hewani ini merupakan kelompok pangan yang bersumber dari kelompok hewani. Ada dua jenis kelompok sumber hewani diantaranya adalah kelompok darat dan kelompok perairan. diantara kelompok bahan hewani darat diantaranya adalah sapi, kerbau, domba, kambing, unggas dan sebagainya. Sedangkan kelompok dari hewan perairan adalah kelompok ikan termasuk di dalamnya adalah ikan, kerang, cumi, udang, kerang dan sebagainya.

2.3. Bahan Pangan Nabati

Bahan pangan nabati menurut Latifah (2020) Bahan pangan nabati adalah bahan pangan yang bersumber dari tumbuhan, seperti buah-buahan, sayur-mayur, kacang-kacangan, dan umbi-umbian.

2.4. Klasifikasi

Menurut Klasifikasi menurut Devita, Dkk (2018) Klasifikasi merupakan cara pengelompokan benda berdasarkan ciri – ciri yang dimiliki oleh objek klasifikasi. Dalam prosesnya, klasifikasi dapat dilakukan dengan banyak cara baik secara manual ataupun dengan bantuan teknologi. Klasifikasi yang dilakukan secara manual adalah klasifikasi yang dilakukan oleh manusia tanpa adanya bantuan dari algoritma cerdas komputer. Sedangkan klasifikasi yang dilakukan dengan bantuan teknologi, memiliki beberapa algoritma, diantaranya *Naive Bayes, Support Vector Machine, Decision Tree, Fuzzy* dan Jaringan Saraf Tiruan. Menurut Putro, Dkk (2020) Secara singkat, klasifikasi merupakan pengelompokan objek kedalam kelas tertentu berdasarkan kelompoknya yang biasanya disebut dengan kelas (*class*). Dari semua pendapat para ahli dapat disimpulkan bahwa klasifikasi adalah proses pengelompokan objek atau benda berdasarkan ciri-ciri yang dimiliki oleh objek tersebut. Proses klasifikasi dapat dilakukan secara manual oleh manusia atau dengan bantuan teknologi menggunakan berbagai algoritma seperti *Naive Bayes, Support Vector Machine, Decision Tree, Fuzzy*, dan Jaringan Saraf Tiruan. Klasifikasi bertujuan untuk mengategorikan objek ke dalam kelas atau kelompok tertentu. Dengan kata lain, klasifikasi membantu kita memahami dan

mengorganisir objek atau data berdasarkan persamaan atau perbedaan karakteristiknya, baik secara manual maupun dengan bantuan komputer.

2.5. RapidMiner

RapidMiner merupakan perangkat lunak yang bersifat terbuka (*open source*). RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap *data mining*, *text mining* dan analisis prediksi. RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik.

2.6. Naive Bayes

Menurut Watratan dan Moeis (2020) Metode *Naive Bayes* merupakan salah satu metode yang menggunakan metode perhitungan probabilitas dan statistik.

a. Kegunaan *Naive Bayes*

1. Mengklasifikasi dokumen teks seperti teks berita ataupun teks akademis;
2. Sebagai metode *machine learning* yang menggunakan probabilitas;
3. Untuk membuat diagnosis medis secara otomatis;
4. Mendeteksi atau menyaring spam.

b. Kelebihan *Naive Bayes*

1. Bisa dipakai untuk data kuantitatif maupun kualitatif.
2. Tidak memerlukan jumlah data yang banyak.
3. Tidak perlu melakukan data *training* yang banyak.
4. Jika ada nilai yang hilang, maka bisa diabaikan dalam perhitungan.
5. Perhitungannya cepat, efisien, mudah dipahami dan mudah dibuat.
6. Pengklasifikasian dokumen bisa dipersonalisasi, disesuaikan dengan kebutuhan setiap orang.
7. Jika digunakan dalam bahasa pemrograman, *code*-nya sederhana.
8. Bisa digunakan untuk klasifikasi masalah *biner* ataupun *multiclass*.

c. Kekurangan *Naive Bayes*

1. Apabila probabilitas kondisional bernilai nol, maka probabilitas prediksi juga akan bernilai nol.
2. Asumsi bahwa masing-masing variabel independen membuat berkurangnya akurasi, karena biasanya ada korelasi antara variabel yang satu dengan variabel yang lain.
3. Keakuratannya tidak bisa diukur menggunakan satu probabilitas saja. Butuh bukti-bukti lain untuk membuktikannya.
4. Untuk membuat keputusan, diperlukan pengetahuan awal atau pengetahuan mengenai masa sebelumnya. Keberhasilannya sangat bergantung pada pengetahuan awal tersebut. Banyak celah yang bisa mengurangi efektivitasnya dirancang untuk mendeteksi katakata saja, tidak bisa berupa gambar

d. Langkah-langkah perhitungan dengan metode *Naive Bayes* sebagai berikut :

1. Menentukan kategori (penyakit) yang muncul berdasarkan data latih.

2. Menghitung nilai probabilitas penyakit dan gejala.
3. Menghitung nilai *bayes* berdasarkan probabilitas penyakit dan gejala yang timbul dan menentukan presentase nilai prediksi kategori.

Menurut Watratan dan Moeis (2020) dalam proses rumus probabilitas *naive bayes* adalah sebagai berikut :

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

Keterangan :

X : Data dengan class yang belum diketahui

C_i : Suatu variabel yang harus dideskripsikan secara probabilistik

P(C_i|X) : Probabilitas hipotesis C_i berdasarkan kondisi X (*posteriori probability*)

P(C_i) : Probabilitas hipotesis C_i (prior probability)

P(X|C_i) : Probabilitas X berdasar kondisi pada hipotesis C_i

P(X) : Probabilitas dari X

Dalam tahap evaluasi hasil, algoritma *naive bayes* diukur kinerjanya dengan cara melakukan pengukuran *precision*, *f1-score*, *recall* dan akurasi. Berikut merupakan rumus dari *f1-score*, *recall* dan akurasi yaitu :

1. Menurut Afif (2020) Akurasi adalah perbandingan jumlah prediksi yang benar. Semua ditentukan dengan mengimplementasikan rumus akurasi berikut:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \times 100 \quad (2)$$

Keterangan :

TP = Prediksi positif yang positif

FN = Prediksi positif yang negative

FP = Prediksi negative yang positif

TN = Prediksi negative yang negative

2. Menurut Afif (2020) Presisi adalah tingkat positif salah adalah perbandingan nilai positif yang salah diklasifikasikan pada kasus *negatif*, yang perhitungannya menggunakan Rumus:

$$Precision = \frac{TP}{TP + FP} \times 100 \quad (3)$$

Keterangan :

TP = Prediksi positif yang positif

FP = Prediksi negative yang positif

3. Menurut Afif (2020) *Sensitivity* disebut juga dengan *recall*. Jika *sensitivity* 100% sama artinya dengan pengklasifikasian menganggap kasus yang diamati positif.

$$Recall = \frac{TP}{TP + FN} \times 100 \quad (4)$$

Keterangan :

TP = Prediksi positif yang positif

FN = Prediksi positif yang negative

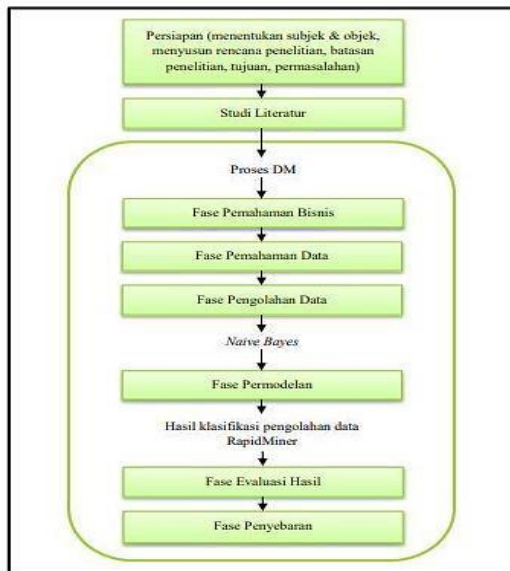
4. Setelah hasil klasifikasi dapat diukur kebenarannya, maka dilakukan perhitungan kombinasi nilai untuk dijadikan sebagai nilai pengukuran (*F1-score*). *F1-score* dapat dihitung dengan rumus sebagai berikut :

$$F1\text{-Score} = 2 \times \frac{Precision \cdot Recall}{Precision + Recall} \quad (5)$$

3. METODE PENELITIAN

3.1. Tahap Penelitian

Pada penelitian ini diambil data penelitian dari data olah pangan serta menggunakan tools *RapidMiner* untuk mengklasifikasi kategori olah pangan pada dinas kesehatan kota palembang menggunakan Metode *Naive Bayes*. Pada bab ini akan membahas metodologi penelitian yang akan digunakan dalam penelitian ini.



Gambar 3.1 Tahapan Penelitian

Dalam gambar 3.1 menjelaskan tahapan penelitian yang akan dilaksanakan oleh peneliti yaitu :

1. Persiapan, tahap ini merupakan tahap penentuan subjek kategori olah pangan pada dinas kesehatan kota palembang di Tahun 2022, objek yang diambil yaitu penentuan klasifikasi kategori olah pangan, batasan dan menyusun rencana penelitian, permasalahan.
2. Tinjauan Kepustakaan, dalam tinjauan kepustakaan dilakukan studi literatur mengenai implementasi data mining untuk penentuan klasifikasi kategori olah pangan menggunakan *Naive bayes*.
3. Fase Pemahaman Bisnis, kegiatan tersebut meliputi tujuan bisnis, penilaian situasi, strategi awal data mining.
4. Fase Pemahaman Data, fase ini melakukan yakni pengumpulan data, mendeskripsikan data dan mengevaluasi kualitas data.
5. Fase Pengolahan Data, fase ini melalui proses *Naive Bayes* dan digunakan data produk pangan sebanyak 730 data yang ada pada Dinas Kesehatan Kota Palembang dalam format dokumen Excel.

6. Fase Permodelan, hasil klasifikasi pengolahan data dengan tool *RapidMiner*.
7. Fase Evaluasi, hasil evaluasi terhadap kualitas data.
8. Fase Penyebaran, hasil penyebaran dari pengolahan data.

3.2. Persiapan dan Pengolahan Data

Pengolahan data mentah pada data produk pangan tahun 2022 dengan melalui proses klasifikasi *naive bayes* untuk memastikan data tersebut layak untuk dijadikan proses pengolahan data melalui klasifikasi data. Pada tahap ini akan dijelaskan secara detail mengenai proses *Preprocessing*. Di bawah ini terdapat gambar 3.2 yang menunjukkan data awal atau data asli dalam format *xlsx* yang akan melalui tahap *Preprocessing*.

1. Data Cleansing

Pada data produk olah pangan diketahui terdapat 7 data yang tidak digunakan dalam penelitian ini yaitu nama pangan, cara penyimpanan, tanggal pengajuan, status oss, nomor hp, nama pelaku usaha dan alamat.

2. Data Selection

Pada data asli produk pangan memiliki Data Transformation 14 variabel. Selanjutnya akan dilakukan seleksi, variabel mana yang akan digunakan dalam proses data mining.

Data semula memiliki 14 variabel, akan tetapi dari 14 variabel tersebut hanya 7 variabel yang digunakan dalam proses data mining antara lain, nama branding produk (CV/PT), jenis pangan, hasil pengolahan pangan, kemasan, ketahanan pangan, bahan pangan dan bentuk pangan. Dari 7 variabel terdapat 2 keterangan, 4 prediktor dan 1 sebagai target yang menjadi kelas *output*.

3. Data Transformasi

Pada tahap ini, dari 7 variabel data yang telah melalui proses seleksi, akan mengalami transformasi data. *Transformasi* data ini melibatkan perubahan tipe data *numerik* menjadi *interval* serta pengurangan nilai atau karakter yang berlebihan pada beberapa variabel.

Adapun penjelasan terkait *transformasi* data pada dataset diatas dapat dilihat pada tabel 3.1 berikut yang merupakan tabel *transformasi* data.

Tabel 3. 1 Transformasi Data

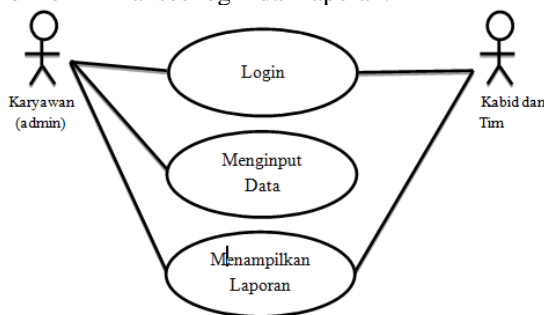
| Variabel | Keterangan | Sebelum Transformasi | Sesudah Transformasi |
|------------------|-------------------------------|--|--|
| Ketahanan Pangan | Jangka waktu ketahanan pangan | Ketahanan pangan dituliskan dengan angka bulan | Ketahanan pangan yang lebih dari satu bulan dituliskan jangka panjang sedangkan dibawah satu bulan ditulis jangka pendek |

| | | | |
|---------------|-------------------------------|---|--|
| Bahan Pangan | Bahan dasar pembuatan makanan | Dilihat melalui variabel nama pangan | Ditentukan melalui nama pangan yaitu berupa hewani dan nabati |
| Bentuk Pangan | Bentuk olahan pangan | Dilihat melalui variabel nama pangan dan jenis pangan | Ditentukan melalui nama pangan dan jenis pangan yaitu berupa bumbu dan non bumbu |

3.3. Rancangan Sistem Olah Produk Pangan

1. Use Case Diagram

Dalam sistem ini aktor yang dimaksud adalah Karyawan, Kabid dan Tim. Karyawan memiliki akses untuk login, input data produk pangan, dan laporan. Kabid memiliki akses login dan laporan.

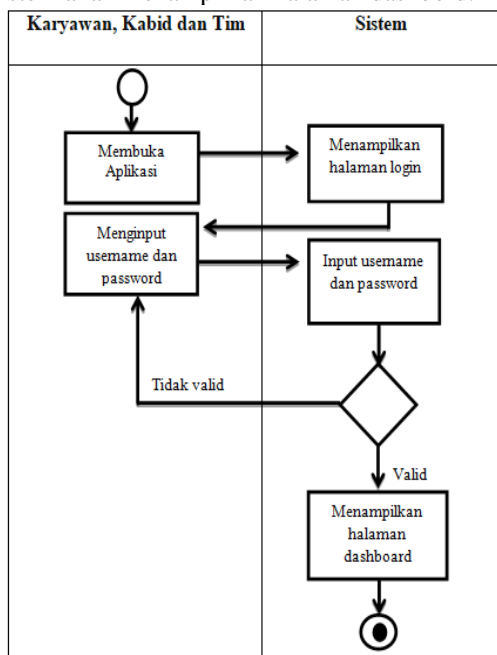


Gambar 3.2 Use Case Diagram

2. Activity Diagram

a. Activity Diagram Login

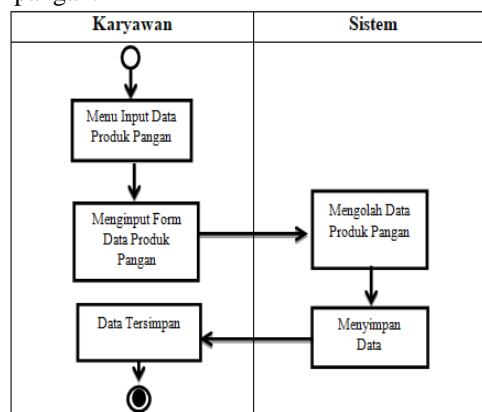
Activity Diagram Login menjelaskan bagaimana Karyawan dan Kabid melakukan proses login. Pertama kali melakukan proses login ke sistem, memasukkan username dan password, sistem akan menampilkan halaman dashboard.



Gambar 3.3 Activity Diagram Login

b. Activity Diagram Input Data

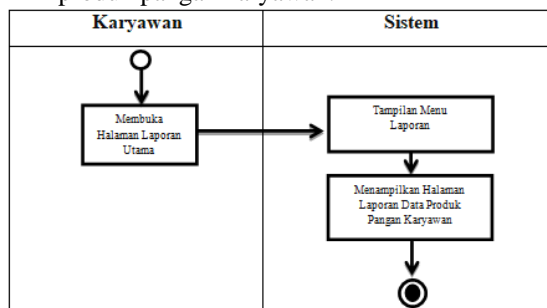
Activity Diagram Input Data menjelaskan bagaimana Karyawan melakukan proses input data produk pangan. Pertama kali sistem menampilkan halaman dashboard menu input data produk pangan setelah itu karyawan memilih menu input data produk pangan lalu sistem akan menampilkan halaman utama menu input data produk pangan, selanjutnya karyawan menginput form data produk pangan lalu sistem akan mengola dan menyimpan data produk pangan.



Gambar 3.4 Activity Diagram Input Data

c. Activity Diagram Laporan Karyawan

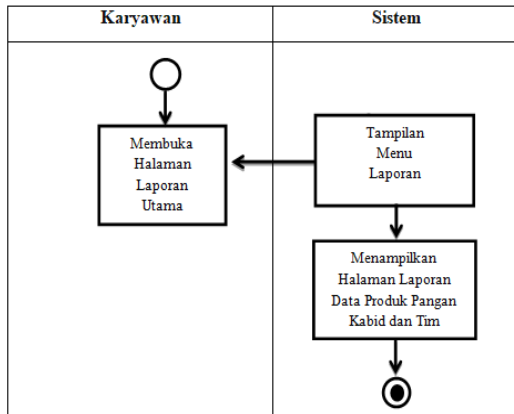
Activity Diagram Laporan menjelaskan bagaimana Karyawan, melakukan proses laporan. Pertama kali sistem akan menampilkan halaman dashboard laporan lalu karyawan dapat memilih laporan data produk pangan karyawan setelah itu sistem akan menampilkan halaman laporan data produk pangan karyawan.



Gambar 3.5 Activity Diagram Laporan Karyawan

d. Activity Diagram Laporan Kabid dan Tim

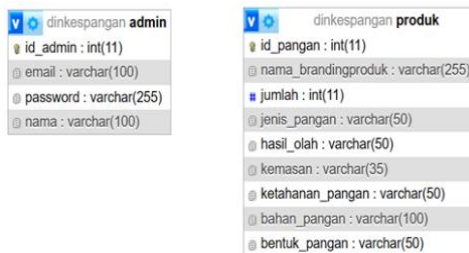
Activity Diagram Laporan menjelaskan bagaimana Kabid dan Tim melakukan proses laporan. Pertama kali sistem akan menampilkan halaman dashboard laporan lalu kabid dan tim dapat memilih laporan data produk pangan kabid dan tim setelah itu sistem akan menampilkan halaman laporan data produk pangan kabid dan tim.



Gambar 3.6 Activity Diagram Laporan Kabid dan Tim

3. Class Diagram

Class Diagram menggambarkan apa saja atribut dan operasi yang ada dalam sistem yang dibuat, class diagram pada website ini terdiri dari Admin dan Data Produk Pangan yang memiliki atribut dan operasinya masing-masing.



Gambar 3.7 Class Diagram

3.4. Tampilan Interface

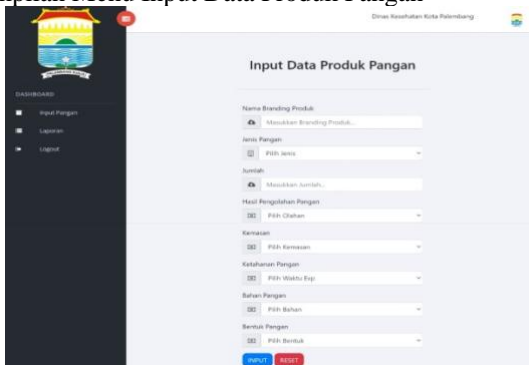
Adapun tampilan *interface* dari penelitian ini sebagai berikut:

1. Tampilan Menu Login



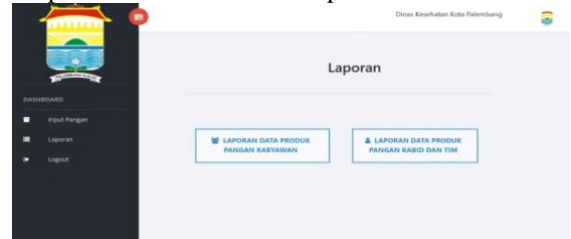
Gambar 3.8 Tampilan Menu Login

2. Tampilan Menu Input Data Produk Pangan



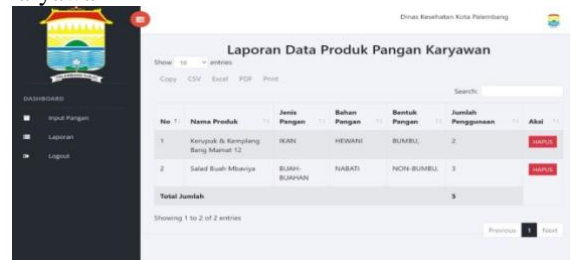
Gambar 3.9 Tampilan Menu Input Data Produk Pangan

3. Tampilan Menu Dashboard Laporan



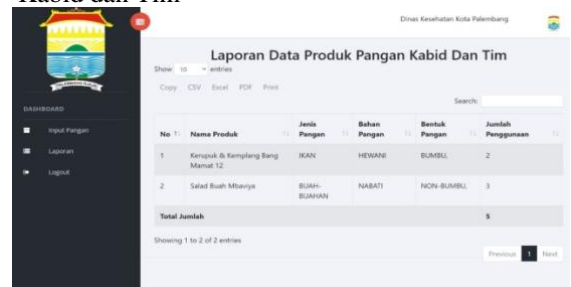
Gambar 3.10 Tampilan Menu Dashboard Laporan

4. Tampilan Menu Laporan Data Produk Pangan Karyawan



Gambar 3.11 Tampilan Laporan Data Produk Pangan Karyawan

5. Tampilan Menu Laporan Data Produk Pangan Kabid dan Tim



Gambar 3.12 Tampilan Menu Laporan Data Produk Pangan Kabid dan Tim

4. HASIL DAN PEMBAHASAN

4.1. Penerapan Metode Naive Bayes

Adapun pengujian terhadap lima variable tersebut, sebagai berikut :

Dik :

$$TP = 42 \text{ (pred.Hewani , true Hewani)}$$

$$FN = 0 \text{ (pred.Hewani , true Nabati)}$$

$$FP = 12 \text{ (pred.Nabati , true Hewani)}$$

$$TN = 165 \text{ (pred.Nabati , true Nabati)}$$

1. Uji Accuracy

Perhitungan dari algoritma *Naive Bayes*, akurasi dilakukan dengan cara jumlah TP + TN dibagi jumlah total data yang diuji. Dengan mengetahui jumlah data yang diklasifikasikan secara benar maka dapat diketahui akurasi hasil prediksi yaitu 94.52%.

Perhitungan Rumus *Accuracy* :

$$\frac{Tp + Tn}{Tp + Fn + Fp + Tn} \times 100$$

$$= \frac{42 + 165}{42 + 0 + 12 + 165} \times 100$$

$$= \frac{207}{219} \times 100$$

$$= 0,9452 \times 100$$

$$= 94,52 \%$$

ConfusionMatrix : *Accuracy*: 94.52% (*positive class*: Nabati)

| | true Hewan | true Nabati | class precision |
|--------------|------------|-------------|-----------------|
| pred Hewan | 42 | 0 | 100.00% |
| pred Nabati | 12 | 165 | 93.22% |
| class recall | 77.78% | 100.00% | |

Gambar 4.1 Hasil *Accuracy*

Hasil pengukuran data *accuracy* yang diperoleh dari data training mencapai 94,52%. Jumlah prediksi hewani yang diklasifikasikan sebagai *true hewani* (merupakan data benar yang terdeteksi benar) oleh *classifier* yaitu 42 data, dan jumlah prediksi hewani yang diklasifikasi sebagai *true nabati* (merupakan data benar namun terdeteksi sebagai data salah) oleh *classifier* yaitu 0, dengan pencapaian *class precision* 100.00%.

2. Uji Precision

Precision adalah jumlah data yang *true positive* (jumlah data *positif* yang dikenali secara benar sebagai *positif*) dibagi dengan jumlah data dikenali sebagai *positif*. Dari hasil pengujian nilai *precision* didapatkan hasil yang berbeda antara hasil prediksi dan hasil perhitungan manual menggunakan metode *naive bayes* dimana didapatkan hasil RapidMiner sebesar yaitu 93.22% untuk *class nabati* sedangkan pada pengujian metode *naive bayes* 78 % untuk *class hewani*.

Perhitungan Rumus *Precision* :

$$\frac{Tp}{Tp + Fp} \times 100$$

$$= \frac{42}{42 + 12} \times 100$$

$$= \frac{42}{54} \times 100$$

$$= 0,78 \times 100$$

$$= 78 \%$$

ConfusionMatrix : *Precision* : 93.22% (*positive class*: Nabati)

| | true Hewan | true Nabati | class precision |
|--------------|------------|-------------|-----------------|
| pred Hewan | 42 | 0 | 100.00% |
| pred Nabati | 12 | 165 | 93.22% |
| class recall | 77.78% | 100.00% | |

Gambar 4.2 Hasil *Precision*

Jumlah prediksi Tidak yang diklasifikasikan sebagai *true hewani* (merupakan data benar yang terdeteksi benar) oleh *classifier* yaitu 12 data & jumlah prediksi Tidak yang diklasifikasikan sebagai *true nabati* (merupakan data benar namun terdeteksi sebagai data salah) oleh *classifier* yaitu 165 data, dengan pencapaian *class precision* 93,22%.

3. Uji Recall

Recall adalah jumlah data yang *true positive* dibagi dengan jumlah data yang sebenarnya *positive* (*true positive* – *true negative*). Untuk nilai *class recall* dengan *true Ya* (merupakan data benar yang terdeteksi benar) mencapai 77.78% untuk *class Hewan* sedangkan untuk *class recall* dengan *true Tidak* (merupakan data benar namun terdeteksi sebagai data salah) mencapai 100% untuk *class Nabati* lalu dilakukan pengujian dengan menggunakan metode *Naive Bayes* pada *RapidMiner* maka didapatkan nilai rata – rata akurasi sebesar 94,52% dengan hasil ini maka penggunaan klasifikasi *naive bayes* untuk prediksi data produk pangan cukup bagus dan dapat dipakai sebagai acuan untuk penelitian selanjutnya.

Perhitungan Rumus *Recall* :

$$\frac{Tp}{Tp + Fn} \times 100$$

$$= \frac{42}{42 + 0} \times 100$$

$$= \frac{42}{42} \times 100$$

$$= 1 \times 100$$

$$= 100\%$$

ConfusionMatrix : *Recall* : 100.00% (*positive class*: Nabati)

| | true Hewan | true Nabati | class precision |
|--------------|------------|-------------|-----------------|
| pred Hewan | 42 | 0 | 100.00% |
| pred Nabati | 12 | 165 | 93.22% |
| class recall | 77.78% | 100.00% | |

Gambar 4.3 Hasil *Recall*

4.4 Uji F1 Score

Perhitungan Rumus F1 Score :

$$\begin{aligned}
 & 2 \times \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \\
 &= 2 \times \frac{78\% \cdot 100\%}{78\% + 100\%} \\
 &= 2 \times \frac{0,78}{1,56} \\
 &= 2 \times 0,5 \\
 &= 1 \\
 &= 100\%
 \end{aligned}$$

4.5 Uji Area Under Curve (AUC)

Tingkat keakuratan kurva ROC/AUC dapat diklasifikasikan menjadi 5 kelompok, yaitu :

1. 0.90 - 1.00 = *Excellent Clasification*
2. 0.80 - 0.90 = *Good Clasification*
3. 0.70 - 0.80 = *Fair Clasification*
4. 0.60 - 0.70 = *Poor Clasification*
5. 0.50 - 0.60 = *Failure*



Gambar 4.4 Diagram Kurva ROC/AUC

Sehingga klasifikasi yang dihasilkan termasuk ke dalam kelompok *Excellent Clasification* karena nilai AUC yang didapatkan dari pengujian berdasarkan kurva ROC menggunakan metode *Naive Bayes* sebesar 0.973.

5. KESIMPULAN

Berdasarkan serangkaian metodologi penelitian yang telah dilaksanakan, penelitian ini menjawab rumusan masalah yang telah di rumuskan pada awal penelitian bahwa pada semua tahapan pengujian diketahui bahan pangan yang lebih banyak digunakan adalah produk olahan bahan pangan nabati dengan keterangan sebagai berikut :

1. Pada uji *Accuracy* didapatkan nilai yang sama antara perhitungan RapidMiner dan perhitungan manual *naive bayes* yaitu sebesar 94,52% *Positive class: Nabati*
2. Pada uji *Precision* didapatkan nilai yang berbeda antara perhitungan rapidminer dan perhitungan manual *naive bayes*, Dimana pada perhitungan RapidMiner di dapatkan hasil yang lebih besar yaitu sebesar 93,22% *Positive class: Nabati* sedangkan di perhitungan manual *naive bayes* di dapatkan nilai sebesar 78%
3. Pada uji *Recall* didapatkan nilai yang sama antara perhitungan RapidMiner dan perhitungan manual

naive bayes yaitu sebesar 100% *Positive class: Nabati*

4. Pada uji *F1 Score* di dapatkan nilai yang sama antara perhitungan RapidMiner dan perhitungan manual *naive bayes* yaitu sebesar 100%
5. Pada uji *Area Under Curve (AUC)* di RapidMiner didapatkan hasil sebesar 0,973 (*Positive class: Nabati*) sehingga berdasarkan tingkat keakuratan kurva ROC/AUC di dapatkan kriteria *Excellent Clasification*.

6. DAFTAR PUSTAKA

- [1] Abdulloh, R. (2018). *7 in 1 Pemrograman Web Untuk Pemula*. Elex Media Komputindo.
- [2] Adi Prayitno, S., & Hartati, F. K. (2020). ILMU DAN PENGETAHUAN BAHAN PANGAN (Bahan Pangan Hewani).
- [3] Afif, A. (2020). Penerapan Algoritma Naive Bayes Untuk Klasifikasi Penyakit Diabetes Mellitus di Rumah Sakit Aisyiah. *JURNAL ILMU KOMPUTER DAN MATEMATIKA*, 1(1), 40-46.
- [4] Anjelita, P., & Rosiska, E. (2019). Rancang Bangun Sistem Informasi E-Learning Pada SMK Negeri 3 Batam. *Computer and Science Industrial Engineering (COMASIE)*, 1(01), 132-141.
- [5] Annur, H. (2018). Klasifikasi Masyarakat miskin menggunakan metode naive bayes. *ILKOM Jurnal Ilmiah*, 10(2), 160-165.
- [6] Anshori, A., Brata, A. H., & Fanani, L. (2023). Pengembangan Sistem Reservasi Rental Kendaraan dan Trip Wisata berbasis Web (Studi Kasus: G19 Tour & Travel). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 7(6), 2729-2735.
- [7] Aprianti, W., & Maliha, U. (2016). Sistem Informasi Kepadatan Penduduk Kelurahan Atau Desa Studi Kasus Pada Kecamatan Bati-Bati Kabupaten Tanah Laut. *Jurnal Sains dan Informatika*, 2(1).
- [8] Astuti, M., MM, M., Matondang, N., Kom, S., & MM, M. T. (2020). *Manajemen Pemasaran: UMKM dan Digital Sosial Media*. Deepublish.
- [9] Asyikin, A. N. (2018). *Pemrograman Web*. Deepublish.
- [10] Devita, R. N., Herwanto, H. W., & Wibawa, A. P. (2018). Perbandingan kinerja metode naive bayes dan k-nearest neighbor untuk klasifikasi artikel berbahasa indonesia. *J. Teknol. Inf. dan Ilmu Komput*, 5(4).
- [11] Doni, R., & Rahman, M. (2020). Sistem monitoring tanaman hidroponik berbasis IoT (Internet of Thing) menggunakan Nodemcu ESP8266. *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, 4(2), 516-522.

- [12] Endra, R. Y., & Aprilita, D. S. (2018). E-Report Berbasis Web Menggunakan Metode Model View Controller Untuk Mengetahui Peningkatan Perkembangan Prestasi Anak Didik. *Explore: Jurnal Sistem Informasi dan Telematika (Telekomunikasi, Multimedia dan Informatika)*, 9(1).
- [13] Fauzi, M. (2019). Pemetaan ketahanan pangan pada badan koordinasi wilayah I Jawa Barat. *Jurnal Industri Pertanian*, 1(1).
- [14] Fitriyyah, S. N. J., Safriadi, N., & Pratama, E. E. (2019). Analisis Sentimen Calon Presiden Indonesia 2019 dari Media Sosial Twitter Menggunakan Metode Naive Bayes. *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, 5(3), 279-285.
- [15] Hadi, I. M., Tukino, T., & Fauzi, A. (2020). Sistem Informasi Monitoring Evaluasi Standar Pembelajaran Menggunakan Framework Codeigniter. In *Conference on Innovation and Application of Science and Technology (CIASTECH)* (Vol. 3, No. 1, pp. 443-452).
- [16] Handayani, P., Susanto, T., Djamaris, A., & Novianti, M. D. (2023). Penerapan Data Mining untuk Menganalisis Faktor-Faktor yang Mempengaruhi Terhadap Kenaikan Harga Minyak Goreng. *Journal of Entrepreneurship, Management and Industry (JEMI)*, 5(4), 251-268.
- [17] Latifah, F. A. (2020). Modul pembelajaran SMA prakarya dan kewirausahaan (pengolahan) Kelas XI: sistem pengolahan makanan khas asli daerah dari bahan pangan nabati dan hewani.
- [18] Lestari, L. (2020). Sistem Pendukung Keputusan Rekomendasi Penerima Bantuan Program Keluarga Harapan (Pkh) Menggunakan Metode Simple Additive Weighting (SAW) (Doctoral dissertation, University of Technology Yogyakarta).
- [19] Lumbangaol, M. H. (2020). Rancang Bangun Sistem Informasi Penjualan dan Penyewaan Properti Berbasis WEB Di Kota Batam (Doctoral dissertation, Prodi Sistem Informasi).
- [20] Maksum, S. R. I., Jamanie, F., & Alaydrus, A. (2019). Strategi Dinas Ketahanan Pangan dalam Meningkatkan Ketahanan Pangan Kota Samarinda. *Pemerintahan Integratif*, 570-581.
- [21] Nugroho, M. (2019). Sistem Informasi Profile Band Online Berbasis Website Menggunakan PHP MYSQL (Studi Kasus: Pernicioushate Band).
- [22] Oscar, B., & Megantara, H. C. (2020). Pengaruh atribut produk terhadap keputusan pembelian produk muslim army. *Pro Mark*, 10(1).
- [23] Pakaya, R., Tapate, A. R., & Suleman, S. (2020). Perancangan aplikasi penjualan hewan ternak untuk qurban dan aqiqah dengan metode Unified Modeling Language (UML). *Jurnal Technopreneur (Jtech)*, 8(1), 31-40.
- [24] Pasi, N., Kadir, A., & Isnaini, I. (2017). Implementasi Sistem Informasi Manajemen Daerah Keuangan Berbasis Akrual pada Pemerintah Kabupaten Dairi. *Jurnal Administrasi Publik (Public Administration Journal)*, 7(1), 49-63.
- [25] Prasetyo, A., & Adhe, N. M. (2018). Pedoman Pembelajaran Algoritma & Pemrograman Dasar. *Tegal, Purbaya E-Journal*.
- [26] Puspita, R., & Widodo, A. (2021). Perbandingan Metode KNN, Decision Tree, dan Naïve Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS. *J. Inform. Univ. Pamulang*, 5(4), 646.
- [27] Rahmasari, T. (2019). Perancangan Sistem Informasi Akuntansi Persediaan Barang Dagang Pada Toserba Selamat Menggunakan Php Dan Mysql. @ *is The Best: Accounting Information Systems and Information Technology Business Enterprise*, 4(1), 411-425.
- [28] Rerung, R. R. (2018). Penerapan data mining dengan memanfaatkan metode association rule untuk promosi produk. *J. Teknol. Rekayasa*, 3(1), 89.
- [29] Sari, A. S., & Hidayat, R. (2022). Designing website vaccine booking system using golang programming language and framework react JS. *JISICOM (Journal of Information System, Informatics and Computing)*, 6(1), 22-39.