PREDIKSI PENGGUNAAN ALAT KONTRASEPSI DI KABUPATEN GORONTALO MENGGUNAKAN METODE *MACHINE LEARNING*

JURNAL

SOFIANI ADIYATRI A. AKILI
NIM : A01420018

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN ILMU KOMPUTER
UNIVERSITAS MUHAMMADIYAH GORONTALO
GORONTALO
2024

Prediksi Penggunaan Alat Kontrasepsi Dikabupaten Gorontalo Menggunakan Metode Machine Learning

Sofiani Adiyatri A. Akili[a] , Wahyudin Hasyim[b], Alter Lasarudin[c],

[abc]Sistem Informasi, Universitas Muhammadiyah Gorontalo. Jl.Prof. Dr. Mansoer Pateda, Desa. Pentadio Timur, Kecamatan. Telaga Biru, Kabupaten.Gorontalo, Provinsi. Gorontalo, Indonesia, Email
Sopianiadiyatri@gmail.com

ABSTRACT

SOFIANI ADIYATRI A. AKILI, Machine Learning techniques are used to predict the use of contraceptive devices in Gorontalo Regency. Wahyudin Hasyim served as supervisor 1. Alter Lasarudin, the second supervisor

This study aims to predict the most populer type of contraceptive using Machine Learning methods, specifically the Principal Component Analyisis (PCA) and Support Vector Machine (SVM) algorithms, based on thw evaluation results of seven types of contraceptives, namely injections, PIL, condoms, PIL, IUD, MOP, and MOW. Gorontalo Regency is one of the areas in indoensia that faces challenges in controlling population growth, despite the continued encouragement of family planning (KB) programs. With an RSME value of 118, the INJECTIVE, contraceptive device achieved the best anticipated RSME value. As a result, the prediction results displayed a very good error value, with a figure ranging from 0.078 to 1.295

Keywords : Support Vector Machine, Principal Component Analysis, Prediction, Contraceptive Devices

ABSTRAK

SOFIANI ADIYATRI A. AKILI, Prediksi Penggunaan Alat Kontrasepsi dikabupaten Gorontalo, menggunakan metode Machine Learning. Dibimbing oleh WAHYUDIN HASYIM sebagai pembimbing 1. Dan ALTER LASARUDIN sebagai pembimbing 2

Indonesia menghadapi tantangan dalam mengendalikan pertumbuhan penduduk, salah satunya di Kabupaten Gorontalo. Meskipun program keluarga berencana (KB) terus didorong, penggunaan alat kontrasepsi di daerah ini dipengaruhi oleh efek samping yang membuat sebagian pengguna merasa tidak nyaman. Penelitian ini bertujuan memprediksi jenis alat kontrasepsi yang paling diminati menggunakan metode Machine Learning, khususnya algoritma Principal Component Analysis (PCA) dan Support Vector Machine (SVM). Berdasarkan hasil evaluasi dari tujuh jenis alat kontrasepsi yaitu Suntik, PIL, Kondom, PIL, IUD, MOP dan MOW. maka nilai RSME yang terbaik hasil prediksinya di dapatkan oleh alat kontrasepsi SUNTIK nilai RMSE 0.118 sehingga hasil prediksi yang dilihat nilai eror yang sangat baik yang menunjukan angka 0.078 – 1.295.

Kata Kunci : Prediksi, Alat Kontrasepsi, Principal Component Analysis, Support Vector Machine

1. PENDAHULUAN

Indonesia, sebagai negara berkembang, menghadapi berbagai tantangan besar, salah satunya adalah masalah kependudukan. Salah satu isu utama adalah tingginya angka pertumbuhan penduduk yang dapat menghambat pembangunan dan kualitas hidup masyarakat. Berdasarkan data Badan Kependudukan dan Keluarga Berencana Nasional (BKKBN), pada tahun 2011, jumlah penduduk Indonesia diperkirakan mencapai 241 juta jiwa, sementara target pemerintah pada tahun 2013 hanya mencapai 243 juta jiwa. Oleh karena itu, BKKBN berupaya mengendalikan angka kelahiran melalui program keluarga berencana (KB), yang mendorong penggunaan alat kontrasepsi, termasuk di Kabupaten Gorontalo.

Di Kabupaten Gorontalo, jumlah Pasangan Usia Subur (PUS) mencapai 811.473 pasangan, dengan lebih dari 555.179 pengguna alat kontrasepsi tercatat pada tahun 2023. BPPKB Gorontalo menyediakan tujuh jenis alat kontrasepsi, yaitu IUD, MOW, MOP, kondom, implan,

suntik, dan pil. Penggunaan alat kontrasepsi ini cenderung meningkat seiring dengan program keluarga berencana yang terus digalakkan. Meskipun alat kontrasepsi banyak tersedia, beberapa faktor seperti efek samping penggunaan menjadi pertimbangan penting bagi masyarakat dalam memilih alat kontrasepsi yang sesuai.

Efek samping seperti siklus haid tidak teratur, perubahan berat badan, mual, sakit kepala, jerawat, dan bintik hitam pada wajah sering kali membuat pengguna merasa tidak nyaman. Oleh karena itu, penting untuk melakukan prediksi terhadap penggunaan alat kontrasepsi yang paling diminati di Kabupaten Gorontalo, berdasarkan faktor-faktor demografis, sosial, dan kesehatan yang memengaruhi keputusan penggunaan. Prediksi ini dapat membantu BPPKB dalam merencanakan program keluarga berencana yang lebih efektif dan relevan dengan kebutuhan masyarakat.

Prediksi adalah salah satu cara untuk meramalkan masa depan dengan mempertimbangkan data historis. Untuk membuat perkiraan, data kuantitatif biasanya digunakan. Prediksi tidak harus memberikan jawaban yang pasti untuk kejadian yang akan terjadi; sebaliknya, mereka harus mencari solusi sedekat mungkin dengan apa yang akan terjadi. Dengan menggunakan data terkini dan historis untuk menilai hasil yang paling mungkin terjadi di masa depan, prediksi bertujuan untuk mengurangi kesalahan dari perbedaan antara hasil yang sebenarnya dan yang diproyeksikan. Prediksi juga bertujuan untuk memperkirakan jenis alat kontrasepsi yang paling banyak diminati oleh masyarakat kabupaten gorontalo. Dengan meganalisis data histori dan faktor-faktor yang mempengaruhi penggunaan, penelitian dapat memberikan gambaran yang lebih akurat tentang alat kontrasepsi ini sendiri.

Metode prediksi menggunakan data historis dapat memberikan gambaran yang lebih akurat tentang jenis alat kontrasepsi yang akan banyak digunakan di masa depan. Dalam hal ini, teknik Machine Learning dapat digunakan untuk menganalisis pola penggunaan alat kontrasepsi yang mungkin tidak dapat ditemukan dengan metode tradisional. Salah satu algoritma yang digunakan dalam penelitian ini adalah Principal Component Analysis (PCA), yang dapat membantu dalam mengidentifikasi faktor-faktor utama yang memengaruhi keputusan masyarakat dalam memilih alat kontrasepsi.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk memprediksi penggunaan alat kontrasepsi di Kabupaten Gorontalo menggunakan metode Machine Learning. Dengan memanfaatkan data yang ada, penelitian ini diharapkan dapat memberikan rekomendasi yang lebih tepat mengenai jenis kontrasepsi yang paling diminati, sehingga BPPKB dapat lebih efektif dalam merancang kebijakan dan program keluarga berencana yang sesuai dengan kebutuhan masyarakat.

2. Tinjauan Pustaka

2.1. Machine Learning

Machine Learning (ML) adalah bidang dalam kecerdasan buatan yang memanfaatkan teknik statistik, probabilitas, dan optimasi untuk mengajarkan komputer mengenali pola dalam data yang besar dan kompleks. Menurut Tom Mitchell (1997), ML memungkinkan komputer untuk mempelajari contoh-contoh sebelumnya dan mengidentifikasi pola yang tidak mudah terlihat dengan metode analisis tradisional.[1]

Arthur Samuel (1950) juga mengemukakan bahwa ML dapat memproses data sebelumnya untuk memprediksi data baru, serta mengeksploitasi pola yang tersembunyi. Ada berbagai model ML yang umum digunakan, seperti klasifikasi, regresi, dan clustering.[2]

2.2. Data Set

Dataset adalah kumpulan data yang disimpan dan memiliki hubungan serta integritas referensi antar tabel, sehingga dapat diproses lebih lanjut untuk berbagai tujuan. Dalam konteks ini, dataset dapat digunakan untuk mengumpulkan informasi tentang penggunaan alat kontrasepsi di Kabupaten Gorontalo. Dengan menggunakan teknik seperti data mining, dataset yang ada bisa diproses untuk menemukan pola dalam perilaku masyarakat terhadap kontrasepsi dan memberikan prediksi mengenai jenis alat kontrasepsi yang paling diminati.[3]

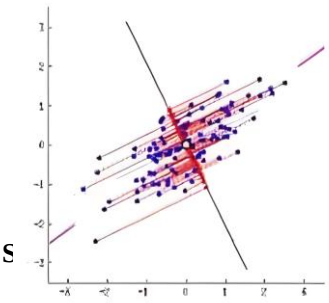
2.3. Prediksi

Prediksi adalah proses memperkirakan kejadian di masa depan berdasarkan data historis, dengan tujuan mengurangi kesalahan antara hasil yang sebenarnya dan yang diprediksi. Prediksi dapat digunakan untuk perencanaan dan pengambilan keputusan dalam banyak bidang, termasuk kesehatan.[4]

2.4. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) adalah teknik statistik yang digunakan untuk menyederhanakan data dengan mengurangi jumlah variabel tanpa mengurangi informasi yang signifikan. PCA dapat membantu dalam mengurangi dimensi data dan mempertahankan variansi yang paling relevan.[5]

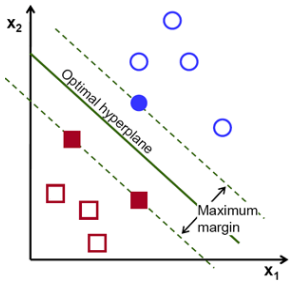
Gambar 1. Principal Component Analisis



2.5. S

Support Vector Machine (SVM) adalah teknik pembelajaran mesin yang digunakan untuk klasifikasi data. SVM bekerja dengan mencari hyperplane yang memisahkan kelas data dengan margin terbesar. Teknik ini telah terbukti efektif untuk klasifikasi data, baik dalam dimensi rendah maupun tinggi, dan digunakan secara luas dalam analisis data kesehatan, termasuk untuk memprediksi preferensi masyarakat terhadap alat kontrasepsi.[6]

Gambar 2. Support Vector Machine



2.6. Penelitian Terkait

Beberapa penelitian sebelumnya telah menggunakan metode Machine Learning dalam prediksi dan analisis data kesehatan juga menggunakan SVM untuk memprediksi harga emas. menggunakan kombinasi antara SVM dan PCA untuk memprediksi penyakit hepatitis, dengan hasil yang menunjukkan akurasi tinggi setelah menggunakan PCA untuk mereduksi dimensi data. Penelitian-penelitian ini menunjukkan potensi penggunaan Machine Learning dalam memproses dan menganalisis data kesehatan, serta memberikan gambaran bagaimana metode ini dapat diterapkan untuk memprediksi penggunaan alat kontrasepsi di Kabupaten Gorontalo.[7]

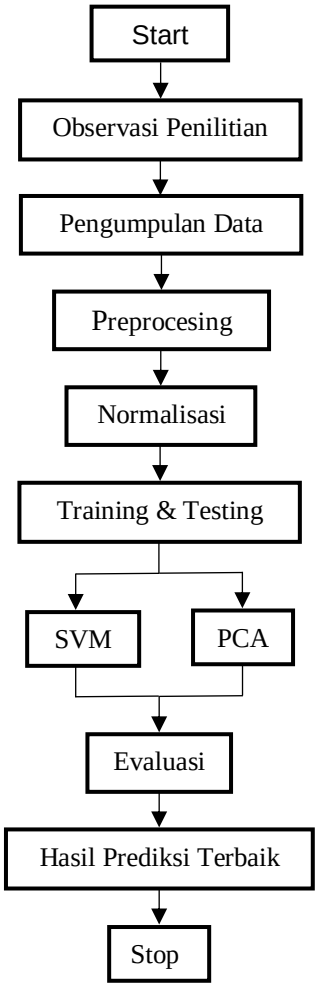
3. Metodologi

3.1. Objek Penelitian

Algoritma *Support Vector Machine* dan *Principal Component Analysis* seperti yang diketahui tidak berdiri sendiri. Penelitian ini dimulai dengan mengumpulkan dataset, dikarenakan penelitian ini bertujuan untuk mengetahui hasil prediksi penggunaan Algoritma SVM dan PCA dalam Machine Learning

3.2 Metode Penelitian

Metode penelitian yang dilakukan memiliki beberapa tahapan seperti *observasi* data, *preprocessing*, *Training* data dan *Testing* data dan penjabaran hasil yang didapatkan. Dalam pengelolaan data ini menggunakan metode *Support Vector Machine* (SVM) dan *Principal Component Analysis* (PCA), pengujian data menggunakan Tool Rapidminer.



3.3 Observasi Penelitian

Observasi penelitian atau yang dimaksud dengan pengumpulan data dimana peneliti mencatat informasi sebagaimana yang mereka lihat selama penelitian

3.4 Dataset

Data set diatas menjelaskan bahwa data penggunaan alat kontrasepsi perbulan ada 228 data untuk tahun 2023. Dalam Dataset tersebut terdapat macam-macam Alat Kontrasepsi seperti Suntik, Pil, Kondom, Implan, IUD, MOP, MOW, MAL.

3.5 Pre-processing Data

Sebelum melakukan Pre-processing data, data diubah bentuk file dari data aktual menjadi data yang dipisah dari 7 alat kontrasepsi tersebut, data dibagi menjadi variabel Date, Alat Kontrasepsi, Label. Perubahan tampilan dilakukan agar olahan data yang menggunakan tools Rapidminer dan perhitungan manual dapat diolah. Selanjutnya dilakukan pre-processing, yang merupakan langkah berikutnya untuk melakukan normalisasi data atau melakukan pre-processing data yang bertujuan untuk mengolah data awal menjadi data yang lebih sederhana atau menghapus variabel yang tidak relevan dari data asli

agar lebih sederhana. Untuk variabel yang dihilangkan adalah Tabel Kabupaten Persen (%), PUS, MAL, Peserta Aktif, Setelah memprocessing data, data diubah menjadi normalisasi data menggunakan tool Rapidminer.

Setelah di Pre-processing, hasil dari data yang dapat diolah dikarenakan tidak relevan, data ini tidak memiliki hubungan signifikan dengan variabel target atau tujuan penelitian. Selanjutnya untuk dataset yang digunakan dalam pengelolaan data adalah sebagai berikut.

Setelah membersihkan data, data diubah bentuk format menjadi atribut, date, Alat kontrasepsi, Label. Data dirangkum perbulan sehingga data yang sebelumnya sebanyak 912 menjadi 54 dataset. Berikut adalah label yang dilakukan transformasi dengan teknik bivot

3.6 Normalisasi Model Min-Max

Model normalisasi min-max membagi rentang nilai, atau nilai maksimum dikurangi nilai minimum dari suatu nilai, kemudian mengurangi nilai minimum dari masing-masing fitur, untuk memungkinkan analisis statistik lebih mudah. Dengan menggunakan rumus min-max sebagai berikut

$$X_i = \frac{0{,}8\ xi - min}{max - min} + 0{,}1$$

Dimana :
xi = Nilai data
X' = Nilai hasil normalisasi
Min(x) = Nilai minimal dari seluruh atribut
Max(x) = Nilai maksimal dari sebuah atribut

3.7 Training Testing

Untuk membuat model Machine Learning, prosedur standar adalah dengan memilih 80% data *Training* dengan jumlah 22 data dan 20% data *Testing* dengan jumlah data 6 dari 28 Dataset Alat Kontrasepsi. Rasio ini memungkinkan model untuk dilatih pada jumlah data yang cukup besar dan kemudian diuji pada jumlah data yang lebih kecil, yang cukup untuk menentukan akurasi model dalam memprediksi hasil nyata. Dengan demikian, model dapat ditingkatkan dan diuji dengan lebih efektif untuk memastikan kinerja yang lebih baik.

3.8 Prediksi Menggunakan PCA

Principal Component Analysis adalah sebuah Teknik statistika untuk mengurangi dimensi dari dataset yang banyak menjadi variabel - variabel yang saling terhubung. Secara khusus Principal Component Analisis melakukan extrasi vaktor-faktor yang memiliki korelasi tinggi dengan mempertahankan karakter aslinya, sehingga pada akhirnya didapatkan variabel yang tidak memiliki korelasi satu sama lainnya.

3.9 Tahapan Penggunaan Algoritma PCA

1. Menghitung mean ($\overline{X}$) dari data pada tiap dimensi menggunakan persamaan

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Dengan :
n = jumlah data sampel
X_i = data sampel

2. Menghitung covariance matrix (C_x) menggunakan persamaan :

$$C_x = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(X_i - \bar{X})^T$$

Dengan :
N = jumlah data sampel
X_i = data sampel
$\bar{X}$ = mean

3. Menghitung eigen vector (v_m) dan eigen value (λ_m) dari covariance matrix menggunakan persamaan :

$$C_x v_m = \lambda_m v_m$$

4. Urutkan eigen value secara descending. Principal Component (PC) adalah deretan eigen vector sesui dengan urutan eigen value
5. Menghasilkan dataset baru.

3.10 Tahapan perhitungan Principal Component Analisis (PCA)

Tahapan perhitungan *Principal Component Analisis* adalah sebagai berikut:

1. Mengambil dataset Alat Kontrasepsi yang telah dibagi menjadi Training dan Testing, data yang digunakan hanyalah data training untuk perhitungan PCA
2. Dataset yang di ambil yaitu 2 alat kontrasepsi Suntik dan pil dan atribut masing-masing memiliki 10 data

   Suntik : (2.12)(-0.45)(2.55)(-0.95)(-0.51)(0.88)(0.09)(-0.16)(0.93)(0.09)

   PIL : (2.63)(-0.01)(0.75)(-1.47)(-0.07)(-0.97)(0.86)(1.04)(0.28)(0.16)
3. Setelah menyiapkan data, hitunglah rata-rata dari 2 atribut data tersebut menggunakan formulasi sebaagi berikut :

$$\text{Var } (A_1) = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{(n-1)}$$

$$\text{Var } (A_2) = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{(n-1)}$$

Sehingga diperoleh nilai rata-rata yaitu A1=(0.46), A2=0.32

4. Selanjutnya menghitung nilai matriks covariance menggunakan formulasi sebagai berikut :

$\begin{pmatrix} Cov\,(A1,A1) & Cov\,(A1,A2) \\ Cov\,(A2,A1) & Cov\,(A2,A2) \end{pmatrix}$=>

$\begin{pmatrix} 1.33 & 0.76 \\ 0.76 & 1.28 \end{pmatrix}$

$\begin{pmatrix} Var\,(A1) & Cov\,(A1,A2) \\ Cov\,(A2,A1 & Var\,(A2) \end{pmatrix}$

|M-λI|= 0 => $\begin{pmatrix} 1.33 & 0.76 \\ 0.76 & 1.28 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 0$

=> $\begin{pmatrix} 1.33 & 0.76 \\ 0.76 & 1.28 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = 0$

=> $\begin{pmatrix} 1.33-\lambda & 0.76 \\ 0.76 & 1.28-\lambda \end{pmatrix} = 0$

= ((1.33- λ)(1.28- λ)) −

$\big((0.76)(0.76)\big) = 0$

= λ^2 – 2.61 λ + 1.1248 = 0

5. Untuk menyelesaikan persamaan kuadrat seperti λ^2 – 2.61 λ + 1.1248 = 0,

anda dapat menggunakan rumus kuadrat. Rumus kuadrat untuk persamaan λ^2 + b λ + c = 0 adalah

$$\lambda = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$$

6. dalam kasus ini, persamaan kuadrat adalah λ^2 – 2.61λ + 1.1248 = 0, sehingga a = 1, b = -2.16 dan c = 1.1248.

1. Hitung diskrimin

Diskriminan D adalah :

D = b^2 – 4ac

D = $(-2.61)^2$ - 4.1.1.1248

D = 6.8121 - 4.4992

D = 2.3129

2. Hitung akar-akar persamaan

Dengan diskriminan D, kita bisa menemukan akar-akar persamaan dengan :

$\lambda = \frac{-b \pm \sqrt{D}}{2a}$

$\lambda = \frac{-(-2.16) \pm \sqrt{2.3129}}{2.1}$

$\lambda = \frac{2.16) \pm \sqrt{2.3129}}{2}$

$\sqrt{2.3129}$ = 1.520

Maka :

$\lambda_1 = \frac{2.61+1.520}{2}$

$\lambda_1 = \frac{4.130}{2}$

λ_1 = 2.065

Dan :

$\lambda_2 = \frac{2.61+1.520}{2}$

$\lambda_2 = \frac{1.090}{2}$

λ_2 = 0.545

7. Setelah menghitung persamaan di atas maka nilai eigen value adalah sebagai berikut :

eigen value

λ_1 = 2.065

λ_2 = 0.545

8. Selanjutnya kita mencari nilai eigen vektor

Mencari nilai menggunakan formulasi di bawah ini

$\begin{pmatrix} 1.33 & 0.76 \\ 0.78 & 1.28 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$

$\begin{pmatrix} 1.33 & 0.76 \\ 0.78 & 1.28 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$= 2.065 $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$

1.33v1 + 0.76v2 = 2.065v1 => -0.735v1 + 0.76v2= 0 v1 = 1.03

0.78v1 + 1.28v2 = 2.065v2 => 0.78v1 - 0.785v2= 0 v2 = 0.96 ((0.96) didapatkan sewenang-wenang)

9. Kita substitusikan nilai v2= 0.90 kedalam persamaan

v1 = $\frac{0.76}{-0.735}$ v2 =1.03401

$\begin{pmatrix} 1.33 & 0.76 \\ 0.78 & 1.28 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$

$\begin{pmatrix} 1.33 & 0.76 \\ 0.78 & 1.28 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$= 0.545 $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$

1.33v1 + 0.76v2 = 0.545 v1 => 0.785v1 + 0.76v1 = 0 v1 = 0.96

0.78v1 + 1.28v2 = 0.545 v2 => 0.78v1 + 0.735v2 = 0 v2 = 1.03

Eigen vector = $\begin{pmatrix} 1.03 & 0.90 \\ 0.96 & 1.03 \end{pmatrix}$

V1 = $\begin{pmatrix} 1.03 \\ 0.96 \end{pmatrix}$ = λ1 = 2.065

V2= $\begin{pmatrix} 0.96 \\ 1.03 \end{pmatrix}$ = λ2 = 0.545

10. Hitunglah total dari hasil varians

Total varians di jelaskan pada komponent utama adalah jumlah nilai eigen :2.065 + 0.545 = 2.610

11. Varians yang dijelaskan oleh masing-masing komponent adalah :

Komponent 1 = $\frac{2.065}{2.610}$ = 0.791 = 79.1%

Komponent 2 = $\frac{2.545}{2.610}$ = 0.209 = 20.9%

Interprestasi hasil : Komponent utama (λ1) : menjelaskan bahwa 79.1% dari total varians ini adalah komponent yang paling signifikan dan dominan dalam data. Untuk Komponent utama (λ2) : menjelaskan 20.9% dari total varians. Ini menunjukan bahwa komponent kedua memberikan kontribusi yang lebih kecil dibandingkan

komponent pertama, namun tetap penting dalam menjelaskan struktur data

3.11. Prediksi Menggunakan SVM Linear

Algoritma *Principal Component Analysis* linear merupakan salah satu algoritma yang paling populer, pelatihan yang dilakukan untuk mengetahui kinerja algoritma Support Vektor Machine linear(Citasi)dalah

Adapun data yang dikumpulkan sebagai contoh adalah sebagai berikut:

Ada 7 fitur dari dataset yaitu Suntik, Pil, Kondom, Implan, IUD, MOP. MOW. Jadi formulasi yang digunakan adalah sebagai berikut. :

$$\frac{1}{2}\|w\|2 = \frac{1}{2}\,(w_2^1 + w_2^2 + w_2^3 \ldots w_2^7$$

Dengan syarat :

$yi(w.x_i + b) > 1, i = 1,2,3,4,5\ldots..n$

$y_i(w_1 . x_1 + w_2 . x_2 + w_3 . x_3 + w_4 . x_4 + w_5 . x_5 + w_6 . x_6 + w_7 . x_7) \geq + 1$ Banyak diminati

$y_1(w_1 . x_1 + w_2 . x_2 + w_3 . x_3 + w_4 . x_4 + w_5 . x_5 + w_6 . x_6 + w_7 . x_7) \geq - 1$ Kurang diminati

4. HASIL DAN PEMBAHASAN

4.1. Pengumpulan Data

Data yang digunakan dalam penelitian adalah data yang di dapatkan setelah turun observasi di Badan Penanggulangan Penduduk dan KB (BPPKB). Data ini mencakup seluruh jenis alat kontrasepsi yang berjumlah 7 jenis dengan jumlah data 228 untuk setiap tahunnya. menggunakan data 5 tahun yaitu dari tahun 2020,2021,2022,2023,2024. Data ini diolah sehingga mendapatkan hasil prediksi penggunaan alat kontrasepsi yang paling banyak diminati oleh Pasangan Usia Subur (PUS).

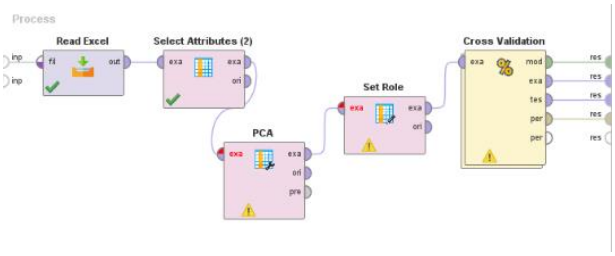
Bab ini memberikan hasil dari perhitungan rumus algoritma *Principal Component Analysis* (PCA) dan *Principal Component Analysis*. Untuk penggunaan algoritma *Principal Component Analisis* (PCA) hanya digunakan untuk mereduksi data seperti mengurangi jumlah variabel dalam data yang kompleks menjadi lebih sedikit. Pelatihan ini diberikan untuk mempelajari kinerja dari *Principal Component Analysis* yang dimaining menggunakan *Principal Component Analisis* (PCA) juga tanpa PCA untuk mendapatkan hasil prediksi yang tepat..

Hasil yang didapatkan berupa 2 macam, yaitu pengelolaan dataset menggunakan algoritma *Principal Component Analysis* yang dimainkan bersamaan dengan *Principal Component Analysis* (PCA) dan juga prediksi penggunaan alat kontrasepsi yang banyak diminati oleh pengguna dengan cara melihat hasil prediksi paling baik di antara 7 alat kontrasepsi. Begitupun untuk dataset yang di mainkan ada berupa dua macam format, yaitu dengan menggunakan keseluruhan alat kontrasepsi (7 Macam) untuk menjalankan algoritma *Principal Component Analysis* dan (*Principal Component Analysis* (PCA), dengan melihat hasil apakah ada perubahan antara *Principal Component Analysis* dan *Principal Component Analysis* (PCA) setelah dimainkan bersamaan atau bisa jadi hanya menggunakan Support Vector Machine tanpa melibatkan algoritma *Principal Component Analysis* (PCA). Kemudian untuk format dataset selanjutnya menggunakan 1 jenis saja, kemudian di lakukan 7 kali sampai semua alat kontrasepsi di maining dan hanya menggunakan algoritma *Principal Component Analysis*. Di karenakan Algoritma *Principal Component Analysis* (PCA) tidak dapat memaining data yang hmenggunakan 1 variabel saja.

4.2 Pengujian Algoritma PCA dan SVM menggunakan Rapidminer

Pengujian menggunakan Software Rapidminer yang sebelumnya seperti yang telah dijelaskan sebelumnya bawasanya tujuannya untuk mengevaluasi efektifitas kombinasi teknik reduksi dimensi dalam meningkatkan performa model prediksi menggunakan alat kontrasepsi menggunakan SVM, serta untuk menganalisis bagaimana kedua algoritma ini dapat membantu dalam mengidentifikasi faktor-faktor utama yang mempengaruhi keputusan pemilihan alat kontrasepsi.

Gambar 3. Proses olah data PCA-SVM



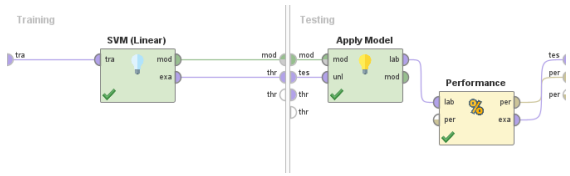
Berdasarkan gambar di atas menunjukan bahwa pengelolaan dataset di rapidminer adalah sebagai berikut :

1. memasukkan operator “Read Excel”, Read Excel sendiri berfungsi untuk menginput / memasukkan data yang akan diolah dirapidminer. Dataset di bagian variabel “alat kontrasepsi” ubah menjadi angka. Angka 1 sampai dengan 7.” Mengikuti jumlah alat kontrasepsi yang ada.
2. Memasukkan operator “Select Attributtes” yang berfungsi untuk memilih atribut / variabel yang akan diolah.
3. Memasukkan Operator “*Principal Component Analysis*” sebagai salah satu algoritma yang digunakan dalam penelitian, fungsi dari *Principal Component Analysis* yaitu untuk mereduksi dimensi

data dengan cara mengubah variabel - variabel asli menjadi satu set variabel baru yang disebut Principal Components.

4. Memasukkan Operator "Set Role", operator ini digunakan untuk memasukkan atribut name dan Target role, yang dimana dalam target role ada beberapa inputan seperti Interpretation, Label, Metadata, Outlier, Prediction, Score, Soure, Weight

Gambar 4. Model Training Testing



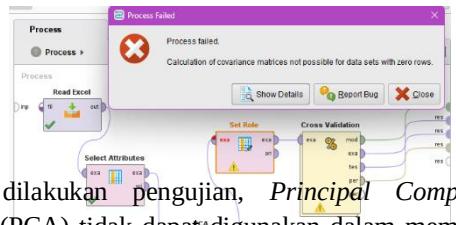
5 Memasukkan Operator " Cross Validation" yang digunakan untuk mengukur kinerja model dengan membagi data menjadi beberapa fold, sehingga model diuji pada data yang berbeda dari data pelatihan.

6 Memasukkan Operator "*Support Vector Machine*", SVM ini adalah algoritma setelah *Principal Component Analysis* sehingga, hasil dari *Support Vector Machine* ini akan menjadi hasil prediksi dengan bantuan operator setelah SVM ini. SVM ini bertujuan untuk mengklasifikasi data dengan mencari hyperplane yang memisahkan kelas-kelas dengan margin terbesar.

7 Memasukkan Operator "Apply Model" ini juga bertujuan untuk menerapkan model yang telah dilatih pada data baru untuk menghasilkan nilai prediksi

8 Memasukkan Operator "Performance" yang digunakan untuk mengevaluasi kinerja model berdasarkan metrik evaluasi yang telah ditentukan.

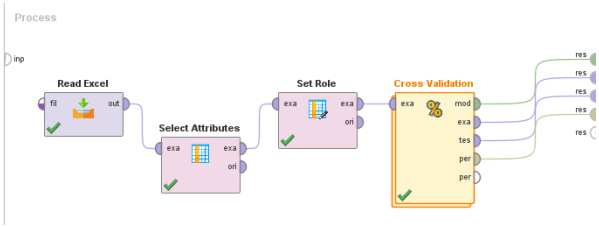
Gambar 5. Hasil PCA



Setelah dilakukan pengujian, *Principal Component Analysis* (PCA) tidak dapat digunakan dalam membantu mereduksi data yang diolah menggunankan *Support Vector Machine* (SVM) dikarenakan algoritma *Principal Component Analysis* (PCA) tidak dapat mengelolah data yang hanya menguji 1 variabel saja. Sehingga pengujian dilakukan Menggunakan 1 algoritma saja yaitu *Support vector Machine* untuk mancari hasil prediksi. Karena pada dasarnya *Principal Component Analysis* (PCA) digunakan untuk mereduksi dimensi data yang memiliki banyak variabel (fitur). Namun, jika hanya memiliki 1 variabel dalam dataset, penerapan *Principal Component Analysis* (PCA) tidak terlalu berguna bahkan tidak dapat diterapkan secara efektif.

4.3 Pengujian Algoritma SVM menggunakan Rapid Miner

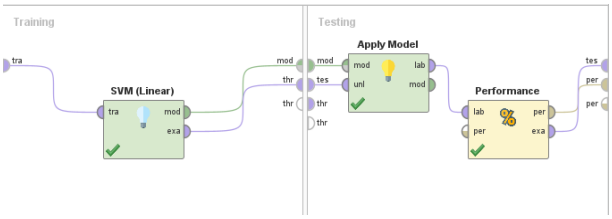
Selanjutnya pengujian rapidminer menggunakan software RapidMiner, operator – operator yang digunakan sebelumnya tetapi ada sedikit penambahan operator.



Gambar 6 Pengujian RapidMiner Menggunakan SVM

Sebelumnya data yang digunakan dalam pengelolaan *Algoritma Principal Component* berbeda dengan data yang dimasukkan kedalam *Support Vector Machine*, karena SVM mengelola data perjenis sesuai banyaknya Alat Kontrasepsi yang diolah untuk melihat hasil akurasi terbaik. Kemudian untuk "*Select Attributes*" sendiri juga memilih variabel apa saja yang akan dimasukkan kedalam pengelolaan dataset tersebut. selanjutnya untuk operator "*Set Role*" digunakan untuk memilih Attribute Name dan Target Role. Untuk Attribute Name memasukkan Label dan Jenis alat kontrasepsi yang tersedia dan target rolenya digunakan untuk Label, Prediction, selanjutnya untuk cross validation juga menggunakan 2-10 kali pengulangan untuk mendapatkan hasil akurasi terbaik.

Gambar 7 Training test untuk SVM Linear



Selanjutnya di dalam operator "*Cross Validation*" terdapat dua bagian terpisah untuk training dan testing. Dalam ruang training terdapat operator *"Nominal to Binominal"* dimana operator tersebut digunakan untuk dataset yang tidak sesuai

dengan yang dapat diolah oleh SVM sehinggal dengan adanya operator tersebut dapat menyesuaikan kebutuhan dari SVM untuk hasil performance. Untuk SVM sendiri adalah algoritma

yang digunakan untuk mengelola data, sehingga data tersebut mendapatkan hasil akurasi terbaik. Kemudian untuk ruang testing terdapat operator "*Apply*

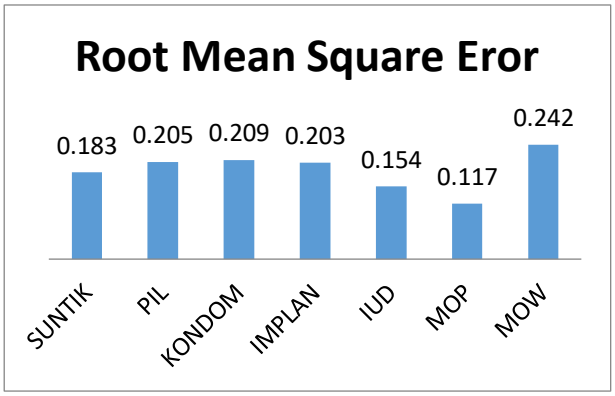
Model" dimana operator tersebut digunakan untuk melihat hasil prediksi dari algoritma SVM. Dan yang terakhir ada operator Performance, dimana operator tersebut digunakan untuk melihat hasil evaluasi seperti *Root Mean Square Eror* (RMSE).

Pengujjian algoritma *Support Vector machine (SVM)* ini menguji 7 jenis alat kontrasepsi tetapi proses pengujiannya menggunakan 1 jenis, sehingga proses pengujian diulangi sampai 7 alat kontrasepsi itu mendapatkan hasilnya prediksinya.

X - Validation	SUNTIK	PIL	KONDOM	IMPLAN	IUD	MOP	MOW
2	0.186	0.223	0.238	0.212	0.158	0.141	0.276
3	0.200	0.211	0.217	0.217	0.159	0.141	0.265
4	0.201	0.226	0.232	0.232	0.156	0.142	0.248
5	0.194	0.205	0.209	0.203	0.154	0.128	0.258
6	0.192	0.220	0.234	0.234	0.159	0.133	0.250
7	0.192	0.212	0.226	0.226	0.156	0.128	0.244
8	0.183	0.205	0.223	0.223	0.154	0.124	0.242
9	0.193	0.215	0.223	0.223	0.155	0.119	0.246
10	0.183	0.207	0.225	0.225	0.156	0.117	0.249

Tabel 2. Proses 10 kali pengulangan

Gambar di atas menunjukan hasil setelah melakukan uji coba Cross Validation untuk mendapatkan hasil terbaik RSME dari 7 atribut alat kontrasepsi, dengan cara 10 kali pengujian.



Gambar diatas menunjukan hasil dari pengelolaan menggunakan dataset Training sehingga mendapatkan hasil RSME terbaik, kemudian di gambarkan menggunakan diagram batang

Setelah menentukan hasil RSME, selanjutnya pengujian dilakukan menggunakan data testing, menggunakan hasil dari Cross Validation Terbaik.

4.4. De-Normalisasi

Denormalisasi dilakukan setelah mendapatkan hasil prediksi dari hasil olahan data dalam rapidminer. De-normalisasi dilakukan menggunakan rumus.

$$X_i \frac{(xp-0.1)(Max-Min)}{0.8} + min$$

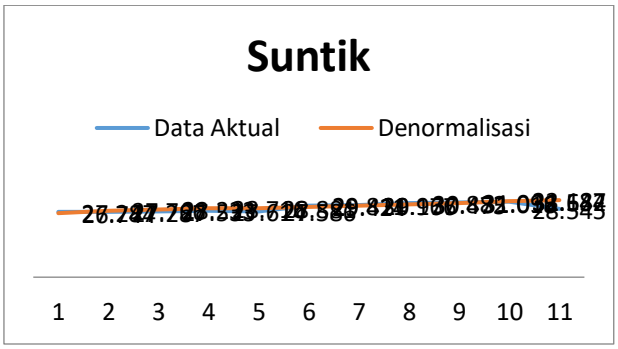
4.5. Hasil Prediksi

Pada hasil prediksi peneliti menggunakan rapidminer dan Microsoft Excel sebagai tools dalam pengelolaan data, rapidminer digunakan untuk mengelola data training sehingga menghasilkan hasil prediksi, yang kemudian dilakukan De-normalisasi data menggunakan Microsoft Excel. Adapun hasil prediksi seperti tabel dibawah ini.
Tabel di bawah adalah tabel hasil prediksi dari penggunaan alat kontrasepsi " Suntik"

Date	Data Aktual (Suntik)	Normalisasi	Prediction	Denormalisasi	Eror	Label
1-Aug-23	27.287	0.2483	0.1699	26.744	-0.543	Sedikit
1-Sep-23	27.287	0.2483	0.3174	27.766	0.479	Sedikit
1-Oct-23	27.533	0.2837	0.3846	28.233	0.700	Banyak
1-Nov-23	27.614	0.2954	0.4544	28.716	1.102	Banyak
1-Dec-23	27.586	0.2914	0.4782	28.881	1.295	Banyak
1-Jan-24	29.811	0.6123	0.5565	29.424	-0.387	Banyak
1-Feb-24	30.176	0.6649	0.6347	29.967	-0.209	Banyak
1-Mar-24	30.882	0.7667	0.7081	30.475	-0.407	Banyak
1-Apr-24	31.096	0.7976	0.7864	31.018	-0.078	Banyak
1-May-24	31.544	0.8622	0.8750	31.632	0.088	Banyak
1-Jun-24	28.545	0.4297	0.9549	32.187	3.642	Banyak

Tabel 3. Hasil Prediksi Suntik

Kemudian didapatkan hasil prediksi dari tabel 18, yang divisualisasikan pada grafik prediksi berikut :

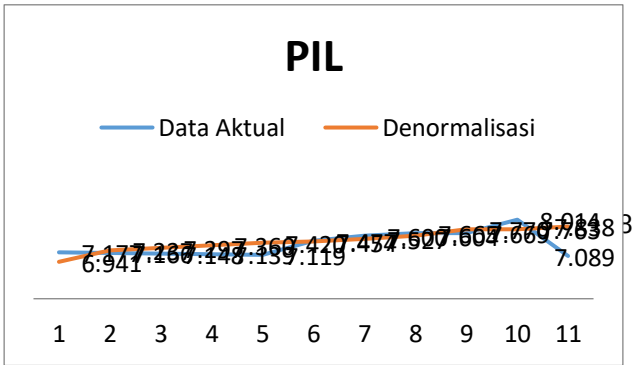


Selanjutnya tabel di bawah adalah tabel hasil prediksi dari penggunaan alat kontrasepsi " PIL" mendapatkan hasil nilai RSME

Date	Data Aktual (PIL)	Normalisasi	Prediction	Denormalisasi	Eror	Label
1-Aug-23	7.177	0.144	0.090	6.941	-0.236	Sedikit
1-Sep-23	7.166	0.142	0.158	7.237	0.071	Sedikit
1-Oct-23	7.148	0.137	0.172	7.297	0.149	Sedikit
1-Nov-23	7.139	0.135	0.186	7.360	0.221	Sedikit
1-Dec-23	7.119	0.131	0.200	7.420	0.301	Sedikit
1-Jan-24	7.477	0.213	0.208	7.454	-0.023	Sedikit
1-Feb-24	7.600	0.241	0.225	7.527	-0.073	Sedikit
1-Mar-24	7.665	0.256	0.242	7.604	-0.061	Sedikit
1-Apr-24	7.669	0.257	0.281	7.770	0.101	Sedikit
1-May-24	8.014	0.337	0.279	7.763	-0.251	Banyak
1-Jun-24	7.089	0.124	0.296	7.838	0.749	Sedikit

Tabel 4. Hasil Prediksi PIL

Kemudian didapatkan hasil prediksi dari tabel 19, yang divisualisasikan pada grafik prediksi berikut :

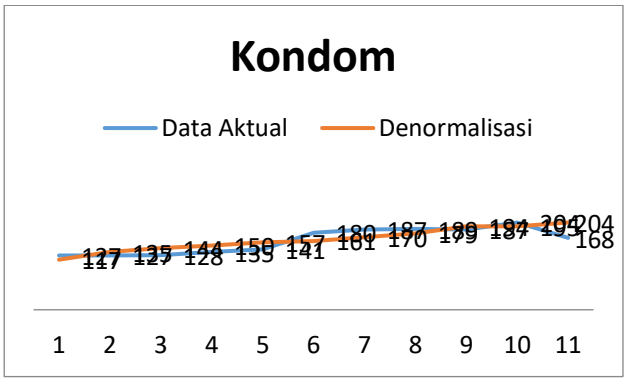


Selanjutnya tabel di bawah adalah tabel hasil prediksi dari penggunaan alat kontrasepsi " Kondom"

Date	Data Aktual (Kondom)	Normalisasi	Prediction	Denormalisasi	Eror	Label
1-Aug-23	127	0.169	0.124	117	-9.823	Sedikit
1-Sep-23	127	0.169	0.208	135	8.453	Sedikit
1-Oct-23	128	0.174	0.247	144	15.957	Sedikit
1-Nov-23	135	0.206	0.273	150	14.623	Banyak
1-Dec-23	141	0.233	0.307	157	16.000	Banyak
1-Jan-24	180	0.413	0.327	161	18.634	Banyak
1-Feb-24	187	0.445	0.367	170	16.901	Banyak
1-Mar-24	189	0.454	0.410	179	-9.623	Banyak
1-Apr-24	187	0.445	0.479	194	7.494	Banyak
1-May-24	204	0.523	0.484	195	8.549	Banyak
1-Jun-24	168	0.357	0.524	204	36.183	Sedikit

Tabel 5. Hasil Prediksi Kondom

Kemudian didapatkan hasil prediksi dari tabel 20, yang divisualisasikan pada grafik prediksi berikut :

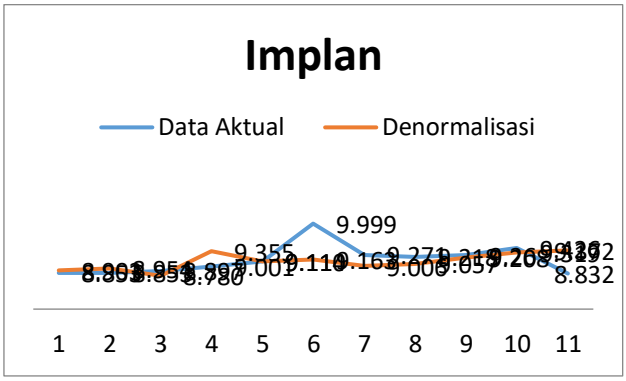


Selanjutnya tabel di bawah adalah tabel hasil prediksi dari penggunaan alat kontrasepsi " Implan"

Date	Data Aktual (Implan)	Normalisasi	Prediction	Denormalisasi	Eror	Label
1-Aug-23	8.853	0.290	0.305	8.903	0.050	Sedikit
1-Sep-23	8.853	0.290	0.321	8.954	0.101	Sedikit
1-Oct-23	8.897	0.303	0.267	8.780	-0.117	Sedikit
1-Nov-23	9.001	0.335	0.444	9.355	0.354	Banyak
1-Dec-23	9.114	0.370	0.369	9.110	-0.004	Banyak
1-Jan-24	9.999	0.642	0.385	9.163	-0.836	Banyak
1-Feb-24	9.271	0.418	0.337	9.006	-0.265	Banyak
1-Mar-24	9.218	0.402	0.352	9.057	-0.162	Banyak
1-Apr-24	9.26	0.415	0.399	9.208	-0.052	Banyak
1-May-24	9.426	0.466	0.433	9.319	-0.107	Banyak
1-Jun-24	8.832	0.283	0.449	9.372	0.540	Sedikit

Tabel 6. Hasil Prediksi Implan

Kemudian didapatkan hasil prediksi dari tabel 21, yang divisualisasikan pada grafik prediksi berikut :

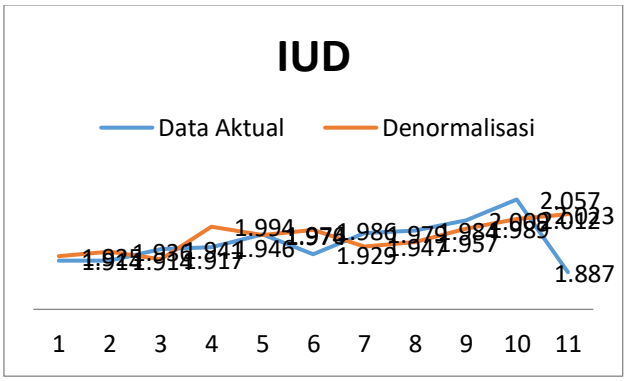


Selanjutnya tabel di bawah adalah tabel hasil prediksi dari penggunaan alat kontrasepsi " IUD"

Date	Data Aktual (IUD)	Normalisasi	Prediction	Denormalisasi	Eror	Label
1-Aug-23	1.914	0.302	0.319	1.925	0.011	Sedikit
1-Sep-23	1.914	0.343	0.335	1.936	0.022	Sedikit
1-Oct-23	1.941	0.302	0.307	1.917	-0.024	Sedikit
1-Nov-23	1.946	0.447	0.424	1.994	0.048	Banyak
1-Dec-23	1.976	0.325	0.393	1.974	-0.002	Banyak
1-Jan-24	1.929	0.401	0.411	1.986	0.057	Banyak
1-Feb-24	1.979	0.351	0.352	1.947	-0.033	Banyak
1-Mar-24	1.984	0.397	0.368	1.957	-0.027	Banyak
1-Apr-24	2.009	0.409	0.417	1.989	-0.020	Banyak
1-May-24	2.057	0.520	0.451	2.012	-0.045	Banyak
1-Jun-24	1.887	0.261	0.468	2.023	0.136	Sedikit

Tabel 8. Hasil Prediksi IUD

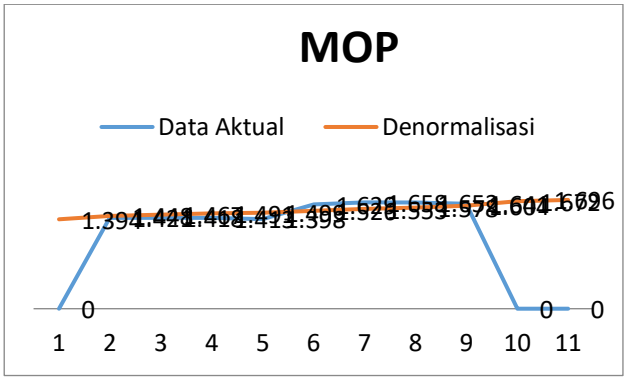
Kemudian didapatkan hasil prediksi dari tabel 22, yang divisualisasikan pada grafik prediksi berikut :



Selanjutnya tabel di bawah adalah tabel hasil prediksi dari penggunaan alat kontrasepsi " MOP"

Date	Data Aktual (MOP)	Normalisasi	Prediction	Denormalisasi	Eror	Label
1-Aug-23	38	0.1220	0.125	38	0.430	Sedikit
1-Sep-23	38	0.1220	0.116	37	-0.754	Sedikit
1-Oct-23	38	0.1220	0.135	40	1.776	Sedikit
1-Nov-23	38	0.1220	0.130	39	1.088	Sedikit
1-Dec-23	39	0.1294	0.131	39	0.231	Sedikit
1-Jan-24	40	0.1367	0.127	39	-1.320	Sedikit
1-Feb-24	41	0.1440	0.132	39	-1.680	Sedikit
1-Mar-24	41	0.1440	0.130	39	-1.910	Sedikit
1-Apr-24	41	0.1440	0.149	42	0.740	Sedikit
1-May-24	40	0.1367	0.153	42	2.246	Sedikit
1-Jun-24	35	0.1000	0.123	38	3.156	Sedikit

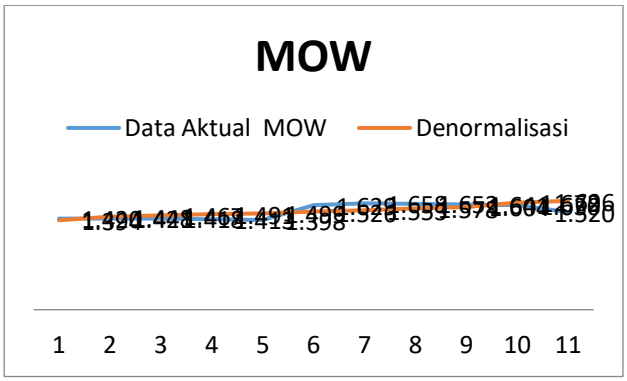
Kemudian didapatkan hasil prediksi dari tabel 23, yang divisualisasikan pada grafik prediksi berikut :



Selanjutnya tabel di bawah adalah tabel hasil prediksi dari penggunaan alat kontrasepsi " MOW"

Date	Data Aktual (MOW)	Normalisasi	Prediction	Denormalisasi	Eror	Label
1-Aug-23	1.420	0.1694	0.120	1.394	-0.026	Banyak
1-Sep-23	1.421	0.1713	0.224	1.448	0.027	Banyak
1-Oct-23	1.418	0.1655	0.260	1.467	0.049	Banyak
1-Nov-23	1.413	0.1559	0.305	1.491	0.078	Banyak
1-Dec-23	1.398	0.1270	0.322	1.499	0.101	Sedikit
1-Jan-24	1.629	0.5723	0.374	1.526	-0.103	Banyak
1-Feb-24	1.658	0.6282	0.425	1.553	-0.105	Banyak
1-Mar-24	1.652	0.6166	0.473	1.578	-0.074	Banyak
1-Apr-24	1.641	0.5954	0.524	1.604	-0.037	Banyak
1-May-24	1.630	0.5742	0.655	1.672	0.042	Banyak
1-Jun-24	1.520	0.3622	0.702	1.696	0.176	Banyak

Kemudian didapatkan hasil prediksi dari tabel 24, yang divisualisasikan pada grafik prediksi berikut :



5. PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan maka, dapat diambil kesimpulannya yaitu :

Berdasarkan hasil penelitian, *Principal Component Analysis* (PCA) tidak dapat menghasilkan nilai RSME. Sehingga pengelolaan data hanya menggunakan algoritma *Support Vector Machine* (SVM), kemudian untuk hasil dari pengelolaah menggunakan *Support Vector Machine* (SVM) dengan menggunakan 7 alat kontrasepsi. Berdasarkan hasil prediksi untuk variabel **Suntik,** menunjukan nilai rata-rata **0.575**. kemudain untuk nilai eror yang dihitung mennunjukan perbedaan antara hasil prediksi dan nilai aktual, dalam hal ini **eror** rata-rata adalah **0.517.** Kemudian hasil prediksi untuk variabel **PIL,** menunjukan nilai rata-rata **0.212**. kemudain untuk nilai eror yang dihitung menunjukan perbedaan antara hasil prediksi dan nilai aktual, dalam hal ini **eror** rata-rata adalah **0.086.** Selanjutnya hasil prediksi untuk variabel **Kondom,** menunjukan nilai rata-rata **0.341** kemudain untuk nilai eror yang dihitung menunjukan perbedaan antara hasil prediksi dan nilai aktual, dalam hal ini **eror** rata-rata adalah **3.198.** Selanjunya untuk hasil prediksi untuk variabel **Implan,** menunjukan nilai rata-rata **0.369**. kemudain untuk nilai eror yang dihitung menunjukan perbedaan antara hasil prediksi dan nilai aktual, dalam hal ini **eror** rata-rata adalah **0.045.** Selanjunya hasil prediksi untuk variabel **IUD,** menunjukan nilai rata-rata **0.386**. kemudain untuk nilai eror yang dihitung menunjukan perbedaan antara hasil prediksi dan nilai aktual, dalam hal ini **eror** rata-rata adalah **0.011.** Selanjutnya hasil prediksi untuk variabel **MOP,** menunjukan nilai rata-rata **0.132** kemudain untuk nilai eror yang dihitung menunjukan perbedaan antara hasil prediksi dan nilai aktual, dalam hal ini **eror** rata-rata adalah **0.364.** Terakhir hasil prediksi untuk variabel **MOW,** menunjukan nilai rata-rata **0.399**. kemudain untuk nilai eror yang dihitung menunjukan perbedaan antara hasil prediksi dan nilai aktual, dalam hal ini **eror** rata-rata adalah **0.012.**

5.2 Saran

Dari hasil penelitian, maka terdapat beberapa saran yang untuk penelitian selanjutnya, terkait data penggunaan alat kontrasepsi, untuk penggunaan Algoritma *Principal Component Analysis*, mereduksi dimensi menggunakan *Principal Component Analysis* (PCA) harus menggunakan lebih dari 1 variabel untuk mendapatkan hasil yang efektif dalam mereduksi data. Kemudian untuk *Support Vector machine*, algoritma SVM terdapat berbagAI jenis dalam Rapidminer, untuk selanjutnya dicoba menggunakan algoritma SVM dengan *Type Kernel* seperti *linaer, polynominial*, atau (RBF). Algortima SVM dapat bekerja menggunakan dataset yang banyak sehingga cocok digunakan.

DAFTAR PUSTAKA

Naibaho, A. F. A., & Zahra, A. (2023). Prediksi Kelulusan Siswa Sekolah Menengah Pertama Menggunakan Machine Learning. *Jurnal Informatika Dan Teknik Elektro Terapan*, *11*(3). https://doi.org/10.23960/jitet.v11i3.3056

Wijaya, Y. F., & Triayudi, A. (2023). Perbandingan Algoritma Klasifikasi Data Mining Pada Prediksi Penyakit Diabetes. *Journal of Computer System and Informatics (JoSYC)*, *5*(1). https://doi.org/10.47065/josyc.v5i1.4614

Kafil, M. (2020). Penerapan Metode K-Nearest Neighbors Untuk Prediksi Penjualan Berbasis Web Pada Boutiq Dealove Bondowoso. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *3*(2), 59–66. https://doi.org/10.36040/jati.v3i2.860

Maiti, & Bidinger. (2020). CLUSTER JUMLAH PENGGUNA ALAT KONTRASEPSI MENGGUNAKAN METODE K-MEANS (Studi Kasus di Kantor KB Kota Yogyakarta). *Journal of Chemical Information and Modeling*, *53*(9), 1689–1699.

Sari, Y. R., Sudewa, A., Lestari, D. A., & Jaya, T. I. (2020). Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer. *CESS (Journal of Computer Engineering, System and Science)*, *5*(2), 192. https://doi.org/10.24114/cess.v5i2.18519

Data, H. S.-J. D., & 2024, U. (2024). Penerapan Algoritma Principal Component Analysis Dalam Reduksi Dimensi Data. *Ilmuteknik.Org*, *1*(5), 1–16. http://www.ilmuteknik.org/index.php/duniadata/article/view/96

Dewi, S. P., Nurwati, N., & Rahayu, E. (2022). Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor. *Building of Informatics, Technology and Science (BITS)*, *3*(4), 639–648. https://doi.org/10.47065/bits.v3i4.1408