

# Penerapan Data Mining Dengan Metode Support Vector Machine Untuk Prediksi Cuaca

Idrianis Fau

Prodi teknik Informatika, Fakultas Ilmu Komputer dan Teknologi informasi, Universitas Budi Darma, Medan, Indonesia  
Email: idrianisfau@gmail.com

**Abstrak**—Cuaca merupakan suatu fenomena alam yang sangat berdampak bagi manusia. Informasi tentang prediksi cuaca sangat dibutuhkan oleh manusia dan terkhusus di Kota Medan. Informasi ini sangat bermanfaat untuk mengetahui kejadian cuaca disekitar kita. Data mining merupakan suatu proses pengumpulan informasi penting pada data yang besar, sehingga dapat membantu untuk mengambil keputusan dan bisa membuat prediksi yang akurat. Sehingga peneliti tertarik untuk melakukan penelitian dalam memprediksi cuaca di Kota Medan dengan menggunakan metode *Support Vector Machine* (SVM) sebagai solusi dalam memprediksi cuaca di Kota Medan. Penerapan data *mining* dengan menggunakan metode SVM membantu menghasilkan akurasi yang tepat untuk cuaca berdasarkan kriteria yang telah ditentukan, metode ini cocok untuk untuk prediksi cuaca karna mampu memberikan penilaian yang jelas dan akurat.

**Kata kunci:** *Cuaca; Prediksi; Data Mining; SVM*

**Abstrak**—Cuaca is a natural phenomenon that greatly impacts humans. Information about weather prediction is needed by humans and especially in Medan City. This information is very useful to know the weather events around us. Data mining is a process of collecting important information on large data, so that it can help to make decisions and can make accurate predictions. So that researchers are interested in conducting research in predicting weather in Medan City by using the Support Vector Machine (SVM) method as a solution in predicting weather in Medan City. The application of data mining using the SVM method helps produce the right accuracy for weather based on predetermined criteria, this method is suitable for weather prediction because it is able to provide a clear and accurate assessment.

**Keywords:** Weather; Prediction; Data Mining; SVM

## 1. PENDAHULUAN

Prediksi cuaca merupakan hal yang sangat penting bagi beberapa aktivitas. Dalam prediksi cuaca, terdapat banyak kondisi yang dapat diamati misalkan suhu udara, kelembaban udara, intensitas matahari, angin, hujan dan kondisi udara lainnya. Prakiraan cuaca dilakukan untuk mengetahui keadaan cuaca yang akan datang berdasarkan kondisi cuaca terakhir dan lampau. Di Indonesia situasi cuaca selalu diumumkan dalam jangka waktu 24 jam melalui prediksi cuaca yang di perbaharui oleh Badan Meteorologi Klimatologi dan Geofisika (BMKG), Departemen Perhubungan. Cuaca sangat erat hubungannya dengan aktivitas manusia, karena faktor cuaca dapat memberikan keterbatasan dalam kegiatan yang dilakukan oleh manusia.

Pemilihan metode yang tepat untuk memprediksi kondisi cuaca ialah kegiatan yang dilakukan oleh para peneliti atmosfer atau cuaca. Berhubung banyaknya permintaan dari berbagai pihak yang membutuhkan informasi tentang atmosfer yang lebih akurat, cepat dan terperinci. Bahkan ada pihak yang menginginkan tersedianya prakiraan atau ramalan cuaca dalam rentang waktu yang cukup kecil misalnya waktu harian, jam, dan bahkan dalam waktu menit. Jadi dengan adanya kebutuhan seperti ini, mendorong peneliti untuk terus melakukan penelitian-penelitian terkait atmosfer cuaca dan juga mengembangkan metode untuk menentukan cuaca sehingga mampu mendapatkan hasil yang baik dan akurat.

Prakiraan dapat dilakukan dengan berbagai macam cara, salah satunya dengan menggunakan metode *data mining*. *Data mining* ialah suatu proses akumulasi data penting yang berukuran besar. Tujuan dari *data mining* ialah memudahkan pengambilan keputusan dan bisa membuat prediksi yang akurat. Kelebihan menggunakan metode *data mining* ialah kajian suatu data yang besar menjadi lebih mudah. Ada pula jenis metode yang termasuk dalam *data mining* yaitu *C4.5*, *Support Vector Machine (SVM)*, *K-Means*.

Berdasarkan penelitian terdahulu mengatakan bahwa peningkatan sistem prediksi hasil panen tanaman padi dengan metode *SVM* dapat dikatakan berhasil dengan nilai akurasi yang tinggi. Terbukti dengan hasil prediksi nilai MAPE sebesar 6635,53, dan nilai RMSE sebesar 1094810,74 [1]. Beberapa contoh metode tersebut, metode terbaik yang bisa digunakan adalah *Support Vector Machine (SVM)*.

*Support Vector Machine (SVM)* ialah merupakan sistem pembelajaran dengan menggunakan ruang hipotesis yang berupa fungsi-fungsi linear di dalam sebuah fitur yang memiliki dimensi tinggi dan dilatih menggunakan algoritma pembelajaran yang berdasarkan teori optimasi. Metode *Support Vector Machine (SVM)* merupakan salah satu teknik yang baru bila dibandingkan dengan teknik lain. Pemilihan fungsi kernel yang tepat dan sesuai merupakan hal yang sangat penting dan diperlukan, sebab fungsi dari kernel tersebut yang akan menentukan *feature space* dimana fungsi dari *klasifier* akan dicari.

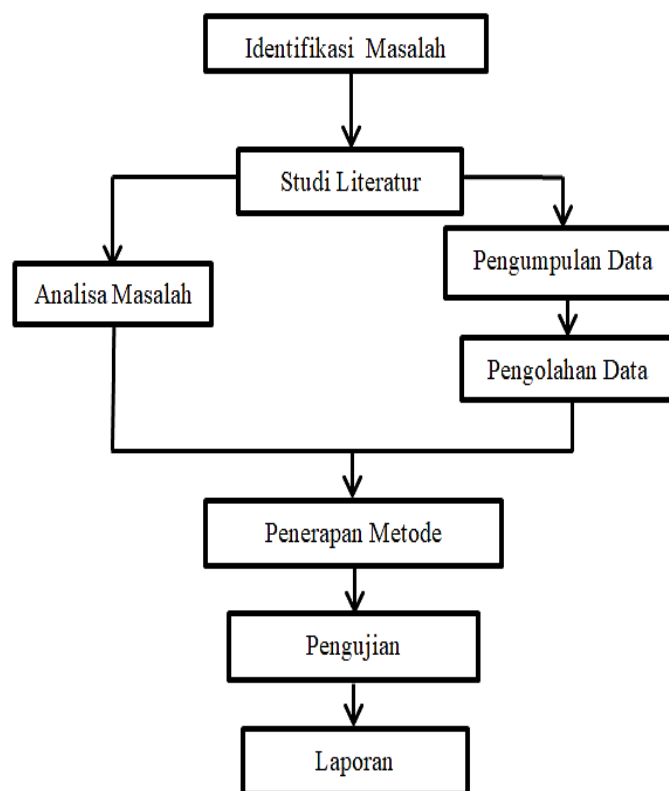
Berdasarkan penelitian terdahulu yang mengimplementasikan metode *SVM* untuk memprediksi siswa yang berpeluang *drop out*, disimpulkan bahwa penerapan metode *SVM* mampu memisahkan siswa yang berpotensi baik dan siswa yang berpeluang *drop out* dengan mengolah kalkulasi data dari nilai akhir, perilaku, dan kehadiran sehingga data tersebut dapat digunakan untuk memprediksi siswa yang *drop out* [2].

Berdasarkan penelitian lainnya yang melakukan prediksi terhadap harga saham PT. Telekomunikasi Indonesia menggunakan metode *SVM* bahwa setelah mengumpulkan data dari 2015-2020 sebanyak 1265 dataset, maka hasil uji pada testing menggunakan metode *SVM* akurasi yang didapat sebesar 0,9641 dengan RMSE 0,0932 sedangkan pada metode *KKN* akurasi yang diperoleh 0,945. Sehingga setelah dilaksanakan perbandingan disimpulkan bahwa metode *SVM* mempunyai tingkat akurasi yang lebih tinggi dibandingkan metode *KKN* [3]. Penelitian lain yang mengimplementasikan metode *SVM* untuk memprediksi lulus tepat waktu pada tugas akhir, dapat disimpulkan bahwa algoritma *SVM* bisa dipergunakan membantu untuk memprediksi kelulusan tugas akhir dengan tepat waktu [4]

Berdasarkan permasalahan dan penjelasan yang telah diuraikan, data mining dengan menggunakan metode *Support Vector Machine* (*SVM*) sangat cocok digunakan untuk prediksi cuaca di Kota Medan, karena metode *SVM* mampu meminimalkan *error* pada *training* set dan memiliki persebaran data yang tidak teratur, sehingga hasil penelitian ini diharapkan bisa menghasilkan model prediksi cuaca kota Medan yang akurat.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian



Gambar 1. Tahapan Penelitian

Penjelasan dari Gambar 1 di atas diantaranya melakukan tahapan Identifikasi Masalah, Tahap pertama yang dilakukan adalah mengidentifikasi masalah yang terjadi mengenai prediksi cuaca dengan menggunakan metode *SVM* untuk mendapatkan solusi dalam permasalahan tersebut. Tahap selanjutnya Studi Literatur, Tahap studi literatur dilakukan dengan mengumpulkan sekelompok data atau informasi yang berkaitan dengan prediksi dan metode *SVM* pada buku ataupun jurnal. Pengumpulan Data, Tahap pengumpulan data dilakukan untuk mengumpulkan sampel data yang akan digunakan dalam penelitian dengan cara mengambil data langsung di website BMKG sesuai arahan pihak BMKG. Tahapan Analisa Masalah merupakan langkah-langkah dari proses suatu penelitian. Analisa masalah yang dilakukan pada tahap ini ialah bagaimana memprediksi cuaca dengan metode *SVM*. Tahap Pengolahan data merupakan rangkaian proses atau cara mengolah data untuk menghasilkan informasi yang berguna sesuai dengan hasil yang diinginkan agar dapat di gunakan. Tahapan Penerapan metode dilakukan sesuai dengan permasalahan yang telah dianalisis. berdasarkan hasil analisis, metode yang diterapkan adalah *Support Vector Machine* (*SVM*). Tahap Pengujian merupakan proses yang dilakukan untuk menilai apakah sistem yang dirancang sudah sesuai dengan yang diharapkan. tahap pengujian ini merupakan kegiatan untuk mengevaluasi keakuratan hasil prediksi, kelebihan dan kekurangan dari sistem. Dan Tahap terakhir yaitu Laporan, pada tahap ini peneliti membuat laporan mengenai hasil penelitian. Laporan yang dibuat didasarkan pada proses penelitian, sehingga laporan tersebut akan menyimpulkan apakah sistem yang dirancang layak untuk digunakan.

## 2.1 Data Mining

*Data mining* ialah suatu proses yang mendapatkan hubungan yang berarti, pola, dan kecenderungan untuk memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika. *Data mining* dibagi atas beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu deskripsi, estimasi, prediksi, klasifikasi, pengklusteran, dan asosiasi [5].

*Data mining* atau *knowledge discovery in database* berarti seluruh proses penggalian atau pengidentifikasian pola, pengetahuan, dan informasi potensial dari kumpulan data besar. Pengetahuan dan informasi yang dihasilkan dari data mining adalah legal, baru, dan mudah dipahami serta berguna [6]. *Data mining* atau *knowledge discovery in database* berarti seluruh proses penggalian atau pengidentifikasian pola, pengetahuan, dan informasi dari kumpulan data yang besar. Data mining dibagi menjadi beberapa kelompok sesuai dengan tugas yang dapat dilakukan yaitu deskripsi, estimasi, prediksi, klasifikasi, clustering, dan asosiasi. Salah satu teknik yang dirancang dalam *data mining* adalah dengan cara menelusuri data yang ada untuk membuat model, kemudian menggunakan model tersebut untuk dapat mengenali pola data lain yang tidak ada dalam database yang tersimpan[7].

Terdapat banyak macam *algoritma* atau metode pencarian data atau informasi dalam data mining. Setiap prosedur penyelesaian memiliki fungsi dan tujuan yang berbeda. Ada beberapa pengelompokan *data mining* [6], yaitu:

1. Deskripsi

Deskripsi bertujuan guna mengidentifikasi pola yang muncul secara berulang pada sebuah data dan mengganti pola tersebut menjadi aturan dan kriteria yang mudah dimengerti oleh para ahli domain aplikasinya.

2. Klasifikasi

Berarti pengelompokan berdasarkan korelasi antara variabel kriteria dan variabel target untuk memberikan identitas suatu data atau objek baru.

3. Prediksi

Prediksi merupakan peranan *data mining* dalam “meramalkan” dimasa yang akan datang nilai kira-kira dari sebuah data.

4. Estimasi

Estimasi adalah peran data mining untuk menghitung nilai perkiraan suatu objek, estimasi hampir sama dengan klasifikasi, perbedaannya terletak pada bentuk pengelompokannya ke arah numerik bukan pada kategorinya.

5. Pengklusteran

*Clustering* berarti mengelompokkan data yang memiliki nilai yang sama (*homogen*). Bentuk data yang dapat dikelompokkan dalam kluster nilai adalah hasil observasi, catatan data, atau kelas data yang memiliki kemiripan.

6. Asosiasi

Asosiasi adalah kumpulan, gabungan, persatuan, atau persekutuan dalam *data mining*. Proses asosiasi adalah pencarian *attribute-attribute* yang ada atau selalu ada pada waktu yang bersamaan.

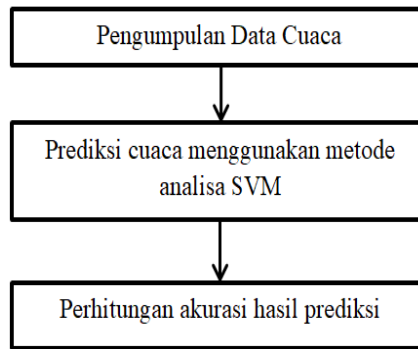
## 2.2 Support Vector Machine (SVM)

*Support vector machine* (SVM) adalah seperangkat metode *supervised learning* untuk klasifikasi, regresi, dan deteksi *outlier*. Kelebihan SVM adalah efektif pada ruang berdimensi tinggi dan menggunakan subset titik pelatihan dalam fungsi keputusan (*support vector*). SVM adalah pengklasifikasi terawasi yang sering digunakan untuk mengekstrak statistik *corpus* untuk analisis sentimen. SVM membentuk *hyperplane* atau satu set *hyperplanes* dalam ruang dimensi tak terbatas yang dapat digunakan untuk klasifikasi atau regresi [10]. Meskipun pelatihan SVM biasanya lambat, namun metode SVM ini sangat akurat karena kemampuannya menangani model nonlinier yang kompleks. SVM kurang rentan terhadap *overfitting* daripada metode lain. SVM dapat digunakan untuk prediksi dan klasifikasi. contoh penerapannya adalah deteksi tulisan tangan, pengenalan objek, identifikasi suara, dan lain-lain [11][12].

## 3. HASIL DAN PEMBAHASAN

Penerapan data mining dalam prediksi berguna untuk menganalisis suatu peristiwa atau kejadian di masa lalu dalam urutan tertentu untuk memprediksi peristiwa yang akan datang. Dan untuk proses prediksinya dilakukan dengan menggunakan metode *Support Vector Machine* (SVM). Prosedur prediksi cuaca pada metode SVM dilakukan dengan terlebih dahulu menyiapkan data-data cuaca yang akan hitung ke dalam metode SVM. Untuk contoh uji metode SVM dilakukan perhitungan sebanyak 10 sampel data yang ada pada Tebal 4.1 dan didapat sebuah model SVM yang bisa memprediksi cuaca di Kota Medan. Proses pengujian ini dilakukan untuk mengetahui akurasi model prediksi dengan metode SVM, seberapa akurat metode SVM mampu memprediksi cuaca di Kota Medan.

Dalam proses pelatihan, ada beberapa variabel yang mampu mempengaruhi rumus seperti nilai alpha, bias, dan weight. Nilai ini memiliki keterkaitan yang erat karena untuk menentukan nilai bias dibutuhkan nilai weight dan nilai alpha. Sedangkan nilai alpha sangat berkaitan dengan nilai yang ada dalam data. Berdasarkan proses ini terbentuk rumus yang memiliki tingkat prediksi yang cukup baik untuk digunakan dalam pengujian data. Adapun proses atau tahapan pengolahan data dapat dilihat pada Gambar di bawah ini:



**Gambar 2.** Tahapan Pengolahan

Berdasarkan Gambar 2, dilakukan pengumpulan data cuaca, tahap ini melakukan pengumpulan data cuaca dari BMKG untuk digunakan pada perhitungan SVM. Prediksi cuaca dengan metode SVM, Tahap ini melakukan prediksi cuaca dan membagi data menjadi data training dan data testing, yaitu data dari bulan januari-april sebagai data training dan data bulan mei sebagai data testing. Perhitungan akurasi klasifikasi hasil prediksi, Perhitungan akurasi klasifikasi hasil prediksi bertujuan untuk mendapatkan hasil akhir dan akurasi data dari data yang dihitung.

Perhitungan metode SVM dalam prediksi cuaca di Kota Medan dapat dilakukan dengan mengambil 10 sampel data dari total keseluruhan data untuk pembuktian cara kerja metode SVM. Dan untuk keseluruhan data akan dihitung menggunakan bahasa pemrograman *Python* pada saat implementasi. Berikut ini tahapan proses SVM :

**Tebal 1.** Contoh Kasus Sampel Data

Tanggal	Temperatur rata-rata (Tavg)	Kelembapan rata-rata (RH_avg)	Curah hujan (RR)	Kecepatan angin rata-rata (ff_avg)	Keterangan
01-01-2022	24.6	96	12.9	2	Hujan
02-01-2022	24.8	94	26.5	2	Berawan
03-01-2022	26.6	85	6.5	1	Hujan
04-01-2022	26.8	88	3.2	1	Hujan
05-01-2022	26.1	84	21.1	2	Hujan
06-01-2022	27.1	84	0.0	2	Berawan
07-01-2022	27.4	80	0.0	1	Berawan
08-01-2022	27.0	87	0.0	1	Berawan
09-01-2022	27.2	84	1.4	2	Berawan
10-01-2022	26.4	84	18.5	2	Berawan
11-01-2022	27.2	84	3.5	1	Berawan
12-01-2022	27.0	78	0.0	1	Berawan

Dua data terakhir pada Tebal 1 digunakan sebagai data testing untuk mendapatkan hasil akurasi dari prediksi cuaca. Ada empat atribut yang digunakan dalam perhitungan ini sesuai dengan parameter yang diamati yaitu Temperatur rata-rata (Tavg), Kelembapan rata-rata (RH\_avg), Curah hujan (RR), dan Kecepatan angin rata-rata (ff\_avg). dan  $w$  memiliki 4 fitur ( $w_1, w_2, w_3,$  dan  $w_4,$ ). Rumus yang digunakan adalah  $y_i (w_1. x_1 + w_2. x_2 + w_3. x_3 + w_4. x_4 \geq + 1$ , sehingga didapatkan nilai masing-masing dari variabel yaitu sebagai berikut :

- $(w_1. 24,6) (w_2. 96) (w_3. 12,9) (w_4. 2) (b) > 1$
- $w_1. 24,8) (w_2. 94) (w_3. 26,5) (w_4. 2) (b) > 1$
- $(w_1. 26,6) (w_2. 85) (w_3. 6,5) (w_4. 1) (b) > 1$
- $(w_1. 26,8) (w_2. 88) (w_3. 3,2) (w_4. 1) (b) > 1$
- $(w_1. 26,1) (w_2. 84) (w_3. 21,1) (w_4. 2) (b) > 1$
- $(w_1. 27,1) (w_2. 84) (w_3. 0,0) (w_4. 2) (b) > 1$
- $(w_1. 27,4) (w_2. 80) (w_3. 0,0) (w_4. 1) (b) > 1$
- $(w_1. 27,0) (w_2. 87) (w_3. 0,0) (w_4. 1) (b) > 1$
- $(w_1. 27,2) (w_2. 84) (w_3. 1,4) (w_4. 2) (b) > 1$
- $(w_1. 26,4) (w_2. 84) (w_3. 18,5) (w_4. 2) (b) > 1$
- $(w_1. 27,2) (w_2. 84) (w_3. 3,5) (w_4. 1) (b) > 1$
- $(w_1. 27,0) (w_2. 78) (w_3. 0,0) (w_4. 1) (b) > 1$

Langkah selanjutnya adalah mencari nilai  $w$  (bobot) dari masing-masing atribut dengan formulasi  $y_i (w_1 \cdot x_1 + w_2 \cdot x_2 + w_3 \cdot x_3 + w_4 \cdot x_4) \geq + 1$ . Untuk mendapatkan nilai dari masing-masing atribut, dilakukan perhitungan linear dengan melakukan eliminasi pada masing-masing variabel.

1. Perhitungan eliminasi variabel 1 dan 2 :

$$\begin{array}{r} 24,6w_1 + 96w_2 + 12,9w_3 + 2w_4 + b = + 1 \\ 24,8w_1 + 94w_2 + 26,5w_3 + 2w_4 + b = + 1 \\ \hline -0,2w_1 + 2w_2 - 13,6w_3 + 2b = 2 \end{array}$$

2. Perhitungan eliminasi variabel 2 dan 3 :

$$\begin{array}{r} 24,8w_1 + 94w_2 + 26,5w_3 + 2w_4 + b = + 1 \\ 26,6w_1 + 85w_2 + 6,5w_3 + 1w_4 + b = + 1 \\ \hline -1,8w_1 + 9w_2 + 20w_3 + 1w_4 + 2b = 2 \end{array}$$

3. Perhitungan eliminasi variabel 3 dan 4 :

$$\begin{array}{r} 26,6w_1 + 85w_2 + 6,5w_3 + 1w_4 + b = + 1 \\ 26,8w_1 + 88w_2 + 3,2w_3 + 1w_4 + b = + 1 \\ \hline -0,2w_1 - 3w_2 + 3,3w_3 + 2b = 2 \end{array}$$

Setelah nilai masing-masing atribut di dapatkan, maka langkah selanjutnya adalah mengeliminasi nilai  $w_1$  (dengan cara perkalian silang indeks  $w_1$ ). Perhitungan eliminasi dari hasil nilai variabel 1 dan 2.

$$\begin{array}{r} -0,2w_1 + 2w_2 - 13,6w_3 + 2b = 2 \\ -1,8w_1 + 9w_2 + 20w_3 + 1w_4 + 2b = 2 \times \\ \hline 0,36w_1 - 3,6w_2 + 24,48w_3 - 3,6b = -3,6 \\ 0,36w_1 - 1,8w_2 - 4w_3 - 0,2w_4 - 0,4b = -0,4 \\ \hline 1,8w_2 + 28,48w_3 - 0,2w_4 - 3,2b = -3,2 \end{array}$$

Setelah nilai  $w_1$  tereliminasi, maka dilakukan perhitungan eliminasi  $w_2$  (dengan cara perkalian silang indeks  $w_2$ ), Perhitungan eliminasi dari hasil nilai variabel 10 dan 11

$$\begin{array}{r} 1,8w_2 + 28,48w_3 - 0,2w_4 - 3,2b = -3,2 \\ -7,2w_2 + 1,94w_3 + 0,2w_4 + 3,2b = 3,2 \times \\ \hline 12,96w_2 - 205,056w_3 + 1,44w_4 + 23,04b = 23,04 \\ 12,96w_2 - 3,492w_3 - 0,36w_4 - 5,76b = -5,76 \\ \hline -201,56w_3 + 1,8w_4 + 28,8b = 28,8 \end{array}$$

Setelah nilai  $w_2$  tereliminasi, maka dilakukan eliminasi  $w_3$  (dengan cara perkalian silang indeks  $w_3$ ), Perhitungan eliminasi dari hasil nilai variabel 18 dan 19

$$\begin{array}{r} -201,56w_3 + 1,8w_4 + 28,8b = 28,8 \\ 6,62w_3 - 1,7w_4 - 6,03b = -6,03 \times \\ \hline -1,334,35w_3 + 11,92w_4 + 190,66b = 28,8 \\ -1,334,35w_3 + 342,65w_4 + 1,215,407b = 1,215,407 \\ \hline -330,73bw_4 - 1,024,48b = -1,024,48 \end{array}$$

Setelah semua bobot di eliminasi, maka langkah selanjutnya mencari nilai setiap atribut  $w$ , sebagai berikut :

1. Mencari nilai  $w_4$

$$0,78w_4 = -0,41$$

$$w_4 = -0,526$$

2. Mencari nilai  $w_3$  dengan menggunakan nilai variabel 24 dengan substitusi  $w_4$  :

$$0,326w_3 - 0,16w_4 = 0,82$$

$$0,326w_3 - 0,16(-0,526) = 0,82$$

$$0,326w_3 + 0,084 = 0,82$$

$$0,326w_3 = 0,82 - 0,084$$

$$0,326w_3 = 0,736$$

$$w_3 = 2,3$$

3. Mencari nilai  $w_2$  dengan menggunakan nilai variabel 17 dengan substitusi  $w_4 = -0,526$ ;  $w_3 = 2,3$  :

$$-1,02w_2 + 1,12w_3 - 0,8w_4 = 2$$

$$-1,02w_2 + 1,12(2,3) - 0,8(-0,526) = 2$$

$$-1,02w_2 + 2,58 + 0,4208 = 2$$

$$-1,02w_2 = 2 - 2,58 - 0,4208$$

$$-1,02w_2 = -1,08$$

$$w_2 = -0,06$$

4. Mencari nilai  $w_1$  dengan menggunakan nilai variabel 8 dengan substitusi  $w_4 = -0,526$ ;  $w_3 = 2,3$ ;  $w_2 = -0,06$  :

$$-0,2w_1 + 3w_2 + 1,4w_3 - 1w_4 = 2$$

$$-0,2w_1 + 3(-0,06) + 1,4(2,3) - 1(-0,526) = 2$$

$$-0,2w_1 - 0,18 + 2,8 + 0,526 = 2$$

$$-0,2w_1 = 2 + 0,18 - 2,8 - 0,526$$

$$-0,2w_1 = -1,146$$

$$w_1 = -0,946$$

Setelah semua nilai setiap atribut  $w$  di dapatkan, maka setelah itu mencari nilai bias ( $b$ ) dengan menggunakan persamaan 10 substitusi  $w_4 = -0,526$ ;  $w_3 = 2,3$ ;  $w_2 = -0,06$ ;  $w_1 = -0,946$  :

$$26,4w_1 + 84w_2 + 18,5w_3 + 2w_4 + b = +1$$

$$26,4(-0,946) + 84(-0,06) + 18,5(2,3) + 2(-0,526) + b = +1$$

$$-214,174 - 5,04 + 42,55 - 1,052 + b = 1$$

$$-177,714 + b = 1$$

$$b = -177,714$$

Setelah semua nilai  $w$  dan nilai  $b$  didapat, maka rangkuman dari datanya yaitu sebagai berikut:

**Tebal 2.** Hasil Nilai  $w$  dan nilai  $b$

Bobot $w$	Nilai $w$	Nilai Bobot ( $b$ )
$w_4$	-0,526	
$w_3$	2,3	
$w_2$	-0,06	-177,714
$w_1$	-0,946	

Setelah nilai  $w$  dan nilai  $b$  di dapatkan, maka langkah selanjutnya adalah melakukan prediksi di data testing yang ada di Tebal sampel data.

$$\text{Kategori } y \begin{cases} +1 \text{ jika } w.z + b > 0 \\ -1 \text{ jika } w.z + b < 0 \end{cases}$$

Keterangan :

+1 jika  $w.z + b > 0$  digunakan untuk syarat mengetahui apakah keterangan dalam suatu data bernilai benar.

-1 jika  $w.z + b < 0$  digunakan untuk syarat mengetahui apakah keterangan dalam suatu data bernilai salah.

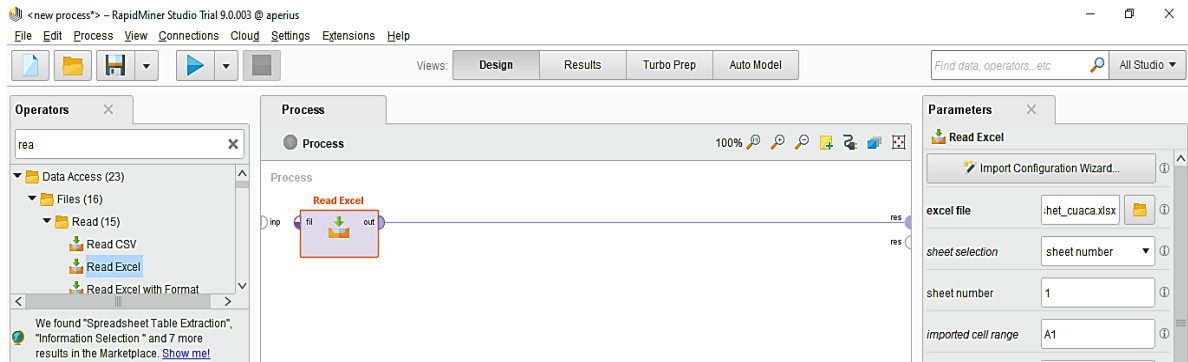
Berikut Uraian untuk mendapatkan hasil akurasi dari data testing :

$$\begin{aligned} 11-01-2022 &= 27,2(-0,946) + 84(-0,06) + 3,5(2,3) + 1(-0,526) - (-177,714) \\ &= -25,731 - 5,04 + 8,05 + 0,526 - (-177,714) \\ &= 155,516 \end{aligned}$$

$$\begin{aligned} 12-01-2022 &= 27,0(-0,946) + 78(-0,06) + 0,0(2,3) + 1(-0,526) - (-177,714) \\ &= -25,542 - 4,68 + 0,526 - (-177,714) \\ &= 148,018 \end{aligned}$$

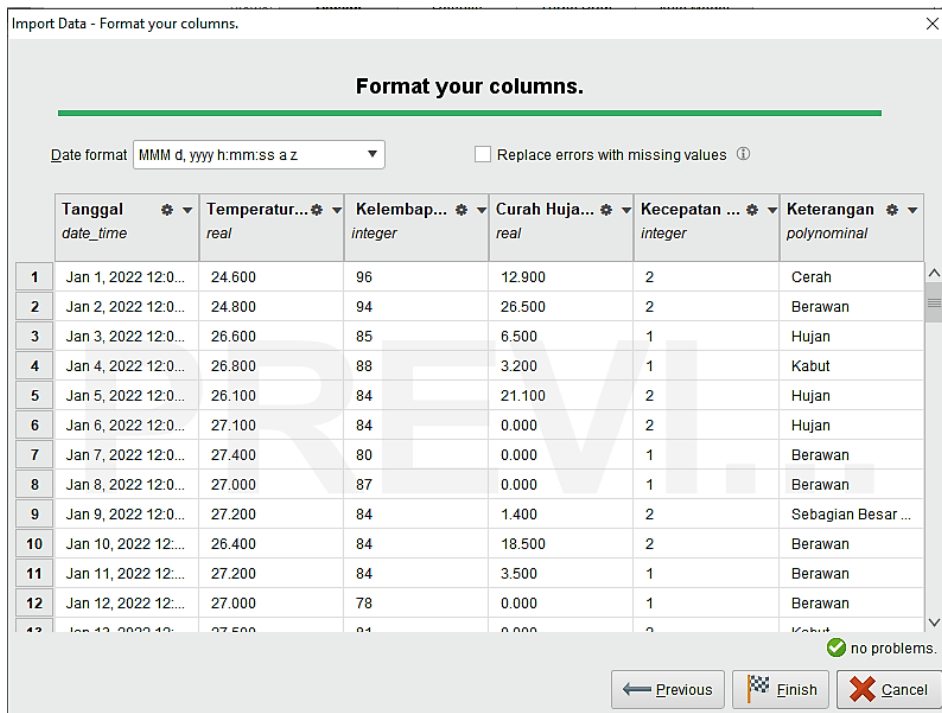
Untuk data 11-01-2022 masuk di kategori yang pertama yaitu di kelas positif +1 jika  $w.z + b > 0$  sehingga data tersebut terprediksi benar (berawan), sedangkan untuk data 12-01-2022 masuk juga di kategori yang sama dengan data sebelumnya yaitu kelas positif, dan datanya terprediksi benar. Jadi, kedua data testing untuk data uji hasil akurasi 100% sesuai dengan data yang ada.

Dalam melakukan proses pengujian pada prediksi cuaca, langkah pertama yang dilakukan adalah *Import Datasheet Excel*. *Import datasheet* cuaca pada *rapiminer* dengan memilih *Read Excel* pada orarator dan meletakkannya pada *main* proses seperti Gambar berikut ini.



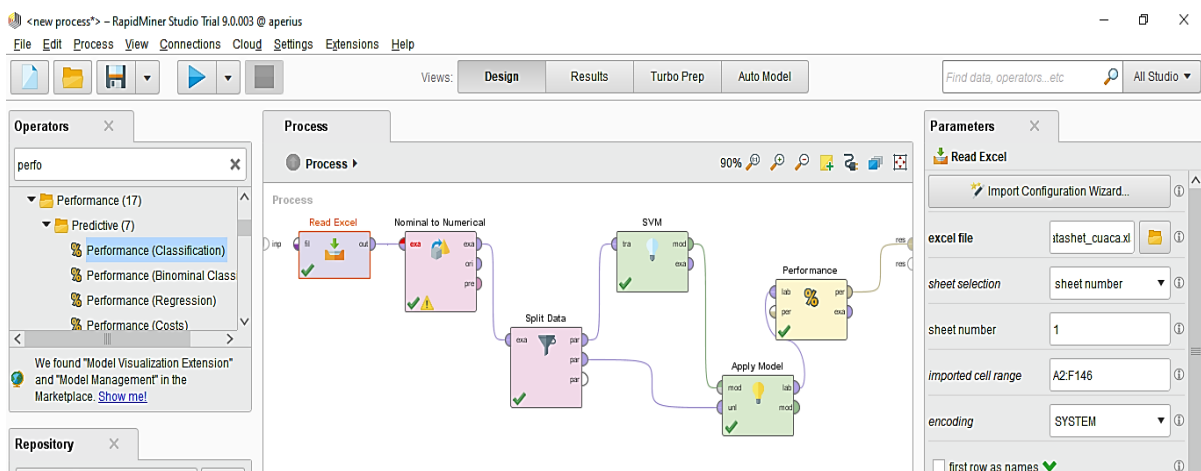
**Gambar 3.** Input Datasheet RapidMiner

Setelah sub menu *Read Excel* berada pada main proses, selanjutnya lakukan impor data dengan memilih menu *import configuration wizard* seperti pada Gambar berikut.



**Gambar 4.** Import Data pada main proses

Melakukan proses prediksi cuaca dengan metode SVM, pada operator pilih, *Support Vector Machine* dan letakkan pada *main* proses posisi didepan *Read Excel* seperti terlihat pada Gambar berikut :



**Gambar 5.** Penerapan Metode SVM

Adapun hasil prediksi cuaca yang diperoleh dari 151 data uji dengan metode SVM pada *rapiminer* seperti Gambar dibawah ini :

Row No.	Tanggal	Keterangan	Temper...	Kelemb...	Curah H...	Kecepa...
1	Jan 1, 20...	Cerah	24.600	96	12.900	2
2	Jan 2, 20...	Berawan	24.800	94	26.500	2
3	Jan 3, 20...	Hujan	26.600	85	6.500	1
4	Jan 4, 20...	Kabut	26.800	88	3.200	1
5	Jan 5, 20...	Hujan	26.100	84	21.100	2
6	Jan 6, 20...	Hujan	27.100	84	0	2
7	Jan 7, 20...	Berawan	27.400	80	0	1
8	Jan 8, 20...	Berawan	27	87	0	1
9	Jan 9, 20...	Sebagian Besar Berawan	27.200	84	1.400	2
10	Jan 10, 2...	Berawan	26.400	84	18.500	2
11	Jan 11, 2...	Berawan	27.200	84	3.500	1
12	Jan 12, 2...	Berawan	27	78	0	1
13	Jan 13, 2...	Kabut	27.500	81	0	2
14	Jan 14, 2...	Kabut	26.700	81	3	1
15	Jan 15, 2...	Kabut	27.100	80	0	2
16	Jan 16, 2...	Sebagian Besar Malam cerah	28.700	78	0	2
150	May 30, ...	Cerah	28.700	81	8888	2
151	May 31, ...	Berawan	28	83	8888	1

**Gambar 6.** Hasil Read Data

accuracy: 54.55%

	true Cerah	true Bera...	true Hujan	true Kabut	true Seb...	true Seb...	true Bad...	true Mala...	true Seb...	true
pred. Ce...	0	0	0	0	0	0	0	0	0	0
pred. Ber...	5	24	1	3	0	1	1	2	6	0
pred. Huj...	0	0	0	0	0	0	0	0	0	0
pred. Ka...	0	0	0	0	0	0	0	0	0	0
pred. Se...	0	0	0	0	0	0	0	0	0	0
pred. Se...	0	0	0	0	0	0	0	0	0	0
pred. Ba...	0	0	0	0	0	0	0	0	0	0
pred. Mal...	0	0	0	0	0	0	0	0	0	0
pred. Se...	0	0	0	0	0	0	0	0	0	0

**Gambar 7.** Hasil Akurasi Prediksi Cuaca

#### 4. KESIMPULAN

Berdasarkan hasil dari penelitian dan pembahasan di bab sebelumnya, maka penulis mengambil suatu kesimpulan terhadap penerapan data mining dengan metode SVM untuk prediksi cuaca di Kota Medan yaitu penerapan data Mining untuk prediksi cuaca di Kota Medan menggunakan *tools* pengujian berupa *RapidMiner* dengan data yang telah ada sesuai dengan data BMKG Medan wilayah 1. Penerapan Metode SVM dengan *rapidMiner* mendapatkan hasil akurasi prediksi cuaca sebesar 54,55%. Melalui penerapan data mining untuk prediksi cuaca memberikan hasil lengkap dan akurat.

#### REFERENCES

- [1] A. S. Budi and P. H. Susilo, "Implementasi Metode Svm Untuk Memprediksi Hasil Panen Tanaman Padi," *Joutica*, vol. 6, no. 1, p. 434, 2021, doi: 10.30736/jti.v6i1.583.
- [2] R. R. Fiska, "Penerapan Teknik Data Mining dengan Metode Support Vector Machine (SVM) untuk Memprediksi Siswa yang Berpeluang Drop Out (Studi Kasus di SMKN 1 Sutera)," *SATIN - Sains dan Teknol. Inf.*, vol. 3, no. 1, p. 15, 2017, doi: 10.33372/stn.v3i1.200.
- [3] W. R. U. Fadilah, D. Agfiannisa, and Y. Azhar, "Analisis Prediksi Harga Saham PT. Telekomunikasi Indonesia Menggunakan Metode Support Vector Machine," *Fountain Informatics J.*, vol. 5, no. 2, p. 45, 2020, doi: 10.21111/fij.v5i2.4449.
- [4] Y. I. Mukti, "Sistem Prediksi Lulus Tepat Waktu Tugas Akhir Mahasiswa Menggunakan Support Vector Machine (Svm),"

- JUTIM (Jurnal Tek. Inform. Musirawas)*, vol. 5, no. 2, pp. 110–115, 2020, doi: 10.32767/jutim.v5i2.1050.
- [5] A. O. P. Dewi, “Big Data di Perpustakaan dengan Memanfaatkan Data Mining,” *Anuva*, vol. 4, no. 2, pp. 223–230, 2020, doi: 10.14710/anuva.4.2.223-230.
- [6] Efori Buulolo, *DATA MINING UNTUK PERGURUAN TINGGI*. Yogyakarta, 2020.
- [7] F. Luis and G. Moncayo, “Konsep Data Mining.”
- [8] L. I. Purba and Dkk., *AGROKLIMATOLOGI*. Yayasan Kita Menulis, 2021.
- [9] M. Pelajaran, *MATERI fungsi atmosfer , pengertian cuaca , iklim , unsur-unsur cuaca dan iklim .* 2021.
- [10] Budiharto Widodo, *PEMROGRAMAN PYTHON UNTUK ILMU KOMPUTER DAN TEKNIK*. ANDI (Anggota IKAPI), 2018.
- [11] R. trias Handayanto and Herlawati, *Data Mining dan Machine Learning Menggunakan Matlab dan Python*. BANDUNG: informatika bandung, 2020.
- [12] T. Imam and Dkk., *Data Mining Dan Data Warehouse Menggunakan Aplikasi KNIME*. 2021.
- [13] I. M. Parapat, M. T. Furqon, and Sutrisno, “Penerapan Metode Support Vector Machine ( SVM ) Pada Klasifikasi Penyimpangan Tumbuh Kembang Anak,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 10, pp. 3163–3169, 2018.
- [14] N. W. Sri wahyuni, “Penerapan Metode Klasifikasi Support Vector Machine (Svm) Untuk Menentukan Karyawan Putus Kontrak Pada Pt. Tae Hang Indonesia,” *J. Inform. SIMANTIK*, vol. 4, no. September, pp. 10–15, 2019.
- [15] A. Roy and A. M. Hendriyawan, “Implementasi Metode Naive Bayes Classifier Untuk Perkitaan Cuaca,” p. 12, 2019.
- [16] A. S. Utami, D. P. Rini, and E. Lestari, “Prediksi Cuaca di Kota Palembang Berbasis Supervised Learning Menggunakan Algoritma K-Nearest Neighbour,” *JUPITER (Jurnal Penelit. Ilmu ...)*, pp. 9–18, 2021.
- [17] S. D. Manullang, E. Buulolo, and I. Lubis, “Implementasi Data Mining Dalam Memprediksi Jumlah Pinjaman Dengan Algoritma C4.5 Pada Kopdit CU Damai Sejahtera,” *J. Sist. Komput. dan Inform.*, vol. 1, no. 3, p. 265, 2020, doi: 10.30865/json.v1i3.2153.
- [18] A. Handayanto, K. Latifa, N. D. Saputro, and R. R. Waliansyah, “Analisis dan Penerapan Algoritma Support Vector Machine (SVM) dalam Data Mining untuk Menunjang Strategi Promosi,” *JUITA J. Inform.*, vol. 7, no. 2, p. 71, 2019, doi: 10.30595/juita.v7i2.4378.
- [19] B. Miftahurrohmah and C. Wulandari, “Analisis Prediksi Mahasiswa Mengundurkan Diri Dari Universitas Xyz Dengan Metode Support Vector Machine,” *Netw. Eng. Res. Oper.*, vol. 4, no. 3, pp. 173–179, 2019.