

Klasifikasi Algoritma *Support Vector Machine (SVM)* Untuk Memprediksi Persebaya Suarabaya Juara BRI Liga 1

Wahyu Kurniawan, Dwi Sukma Donoriyanto

Program Studi Teknik Industry, Fakultas Teknik

Universitas Pembangunan Nasional Veteran Jawa Timur

21032010056@student.upnjatim.ac.id, Dwisukama.ti@upnjatim.ac.id

Alamat : Jl. Rungkut Madya No.1, Gn. Anyar, Kec. Gn. Anyar, Surabaya, Jawa Timur 60294

Korespodensi email : 21032010056@student.upnjatim.ac.id

Abstract. *This research uses the Support Vector Machine (SVM) algorithm to predict Persebaya Surabaya's ranking in BRI Liga 1. The data used includes goals scored, goals given away, total end-of-season points, and status as champions. The results of the analysis using Orange software show that Persebaya Surabaya does not necessarily become a champion if it has a point value of 42 and an SVM value of 41. To become a champion, Persebaya Surabaya must score 69 points or more in a season and achieve an average of more than 54 goals per season. The suggestion of this research is to have more data so that the results of data processing using Orange software are more optimal and accuracy is more precise.*

Keywords: *Clustering, Support Vector Machine, Orange*

Abstrak. Penelitian ini menggunakan algoritma *Support Vector Machine (SVM)* untuk memprediksi peringkat Persebaya Surabaya dalam BRI Liga 1. Data yang digunakan meliputi gol yang dicetak, gol yang diberikan, total poin akhir musim, dan status sebagai juara. Hasil analisis menggunakan *software Orange* menunjukkan bahwa Persebaya Surabaya tidak serta merta menjadi juara jika memiliki nilai poin 42 dan nilai SVM 41. Untuk menjadi juara, Persebaya Surabaya harus mencetak 69 poin atau lebih dalam satu musim dan mencapai rata-rata lebih dari 54 gol per musim. Saran penelitian ini adalah untuk memiliki data yang lebih banyak agar hasil pengolahan data menggunakan *software Orange* lebih optimal dan akurasi lebih tepat.

Kata Kunci: *Klasemen, Support Vector Machine, Orange*

PENDAHULUAN

Bagi yang menempati peringkat pertama, hasil pertandingan akan disimpan di tabel klasemen untuk menentukan juara di akhir musim. Dari segi perkiraan, hasil klasemen setiap tahun dapat dikategorikan berdasarkan sejarah kinerja klub. Salah satu cara untuk mengantisipasi atau menarik kesimpulan dari hasil klasifikasi ini adalah dengan mengklasifikasikan kemungkinan suatu tim memenangkan liga berdasarkan klasemen liga di akhir musim. Temuan ini memberikan sebuah tolok ukur: klub yang ingin memenangkan gelar harus mampu memperoleh tingkat nilai tertentu.

Hasil klasemen BRI Liga 1 tahun 2024 akan dikategorikan oleh SVM. Informasi yang diproses akan Urutkan gol yang dicetak musim ini, gol yang diberikan musim ini, total poin akhir musim, dan apakah tim tersebut juara atau tidak, yang akan ditentukan dengan menggunakan sampel tim Manchester United. Dengan menggunakan klasifikasi ini, jumlah ciri yang kemungkinan besar akan mempengaruhi peluang tim untuk memenangkan liga setiap

musim dapat ditentukan. Penelitian ini memanfaatkan program Rapid Miner 8.1 untuk memudahkan klasifikasi data.

METODOLOGI PENELITIAN

II.1 Data Mining

Penambangan data mengacu pada serangkaian prosedur yang digunakan untuk mengekstraksi nilai lebih dari sekumpulan data dalam bentuk pengetahuan yang tidak dapat ditemukan melalui cara manual. Akibatnya, penambangan data adalah teknik penting untuk mengubah data menjadi informasi. *Apriori, Page Rank, AdaBoost, K-Nearest Neighbors (KNN), Naïve Bayes, Classification and Regression Tree (CART), K-Means, Support Vector Machines (SVM), C4.5, dan Expectation Maximization Algorithm (EM)* adalah sepuluh (10) algoritma penambangan data terbaik, menurut *IEEE International Conference on Data Mining (ICDM)*, salah satu konferensi internasional terbesar mengenai masalah ini. Tidak disebutkan algoritma terbaik di antara sepuluh algoritma tersebut. Oleh karena itu, hal ini menjadi salah satu motivasi dilakukannya penelitian. Salah satu pendekatan yang dapat diterapkan adalah algoritma C4.5. untuk membuat pohon keputusan. Pohon keputusan merupakan suatu teknik yang sangat sederhana untuk dipahami orang. Pohon keputusan adalah model prediksi hierarki atau terstruktur pohon. Mengubah data menjadi pohon keputusan dan aturan keputusan adalah ide di balik pohon keputusan. Keuntungan utamanya adalah menyederhanakan prosedur pengambilan keputusan yang rumit sehingga pengambil keputusan dapat lebih cepat mengidentifikasi solusi untuk permasalahan saat ini. (Nasrullah, A. H. 2021).

II.2 Klasifikasi

Proses penambangan data yang disebut klasifikasi mengatur atau mengelompokkan data menurut kriteria yang telah ditentukan. Ini adalah teknik pembelajaran terawasi yang menciptakan aturan untuk mengelompokkan data pengujian ke dalam kelas atau kelompok yang telah ditentukan sebelumnya menggunakan data pelatihan berlabel. Istilah "metode klasifikasi" mengacu pada proses pengelompokan data dalam gudang data yang diselidiki menggunakan algoritma yang telah ditetapkan. Pendekatan ini berguna untuk operasional perusahaan seperti pemasaran dan penjualan yang memerlukan informasi kategori. Ia mampu memanfaatkan sejumlah metode, termasuk pembelajaran pohon keputusan dan tetangga terdekat. Selain berguna untuk eksplorasi data, pohon keputusan dapat mengungkapkan korelasi tersembunyi antara variabel tujuan dan beberapa variabel masukan potensial. Pohon keputusan berfungsi dengan baik sebagai tahap awal dalam eksplorasi dan pemodelan data

karena pohon keputusan berfungsi dengan baik sebagai tahap awal dalam eksplorasi dan pemodelan data bahkan ketika ia berfungsi sebagai model pamungkas untuk metode lain, yaitu proses pemodelan (Lubis, A. 2020).

II.3 *Support Vector Machine*

SVM adalah sekelompok teknik pembelajaran terbimbing terkait yang digunakan dalam diagnostik regresi dan klasifikasi dalam kedokteran. Secara bersamaan, SVM memaksimalkan margin geometris dan meminimalkan kesalahan klasifikasi empiris. Jadi, Pengklasifikasi Margin Maksimum (SVM) adalah namanya. SVM adalah metode umum yang mengandalkan apa yang disebut gagasan minimalisasi risiko struktural, yang didasarkan pada batasan risiko yang terjamin dari teori pembelajaran statistik. Fraksi maksimum poin dalam kelas yang sama pada sebuah bidang ditemukan oleh SVM di sebuah hyperplane. Salah satu alat untuk membedakan fitur adalah fungsi *hyperplane*. Garis adalah fungsi yang digunakan untuk mengklasifikasikan fitur dalam dua dimensi; bidang adalah fungsi yang digunakan untuk mengklasifikasikan fitur dalam tiga dimensi; *hyperplane* adalah fungsi yang digunakan untuk mengklasifikasikan titik-titik dalam dimensi yang lebih tinggi. Tetap saja, itu Hasil akurasi yang diperoleh masih belum cukup untuk tergolong baik, dan belum cukup digeneralisasikan. Akibatnya, fungsi kernel—yang meningkatkan akurasi ke dimensi yang lebih tinggi—harus digunakan. Suatu fungsi k disebut fungsi kernel jika dapat memenuhi batasan $(x, z) = \phi(x) \cdot \phi(z)$, dimana $\phi(\cdot)$ menyatakan fungsi pemetaan dari ruang masukan ke ruang fitur untuk semua masukan vektor x, z . Dengan kata lain, fungsi kernel adalah fungsi produk dalam ruang fitur. Tanpa mendefinisikan fungsi pemetaan dari ruang masukan ke ruang fitur, model dapat diimplementasikan dalam ruang berdimensi lebih tinggi (ruang fitur) berkat fungsi kernel. Dengan demikian, hyperplane dapat secara efektif berfungsi sebagai batasan keputusan. (Umbara, F. R. 2022).

HASIL DAN PEMBAHASAN

1.1 *Preprocessing Data*

Data cleaning diterapkan untuk menambah isi atribut yang hilang atau kosong, dan merubah data yang tidak konsisten.

1. Data Transformasi Dalam proses ini, data ditransformasikan ke dalam bentuk yang sesuai untuk proses data mining.
2. Data reduksi data dilakukan dengan menghilangkan atribut yang tidak diperlukan sehingga ukuran dari database menjadi kecil dan hanya menyertakan atribut yang

**KLASIFIKASI ALGORITMA SUPPORT VECTOR MACHINE (SVM) UNTUK MEMPREDIKSI PERSEBAYA
SUARABAYA JUARA BRI LIGA 1**

diperlukan dalam proses data mining, karena akan lebih efisien terhadap data yang lebih kecil. Berikut ini adalah tabel data set yang akan digunakan.

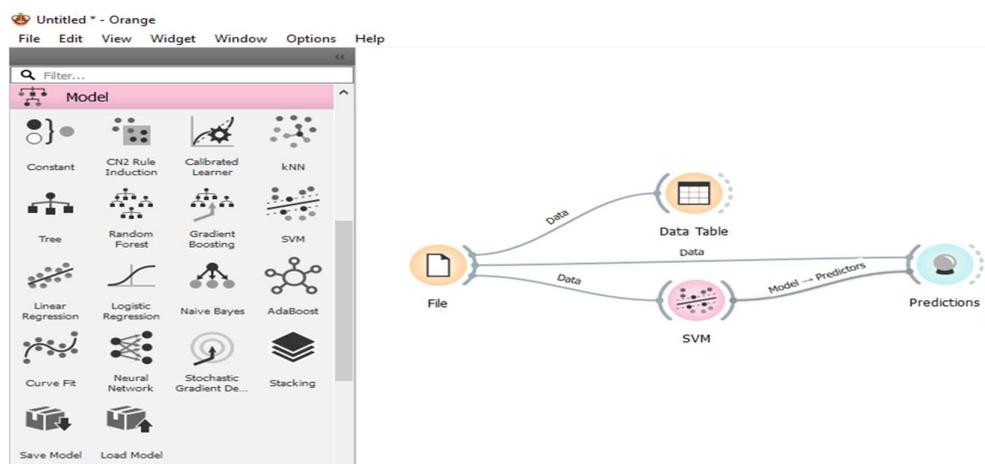
Tabel 3.1 Data set

No	TEAM	PERTANDINGAN	WIN	DRAW	LOS	GF	GA	GD	POIN
1	Borneo FC	34	21	7	6	54	31	23	70
2	Persib Bandung	34	16	14	4	65	38	27	62
3	Bali United FC	34	17	7	10	57	41	16	58
4	Madura United	34	15	10	9	58	45	13	55
5	Dewa United FC	34	14	12	8	59	48	11	54
6	PSIS Semarang	34	15	8	11	49	41	8	53
7	Persis Solo	34	14	8	12	50	47	3	50
8	Persija Jakarta	34	12	12	10	49	41	8	48
9	Persik Kediri	34	13	9	12	58	55	3	48
10	PS Barito Putera	34	11	13	10	51	48	3	46
11	PSM Makassar	34	11	11	12	44	39	5	44
12	Persebaya Surabaya	34	10	12	12	33	46	-13	42
13	PSS Sleman	34	9	12	13	49	53	-4	39
14	Persita Tangerang	34	10	9	15	44	63	-19	39
15	Arema FC	34	10	8	16	42	60	-18	38
16	Rans Nusantara FC	34	8	11	15	36	52	-16	35
17	Bhayangkara FC	34	5	11	18	42	57	-15	26

No	TEAM	PERTANDINGAN	WIN	DRAW	LOS	GF	GA	GD	POIN
18	Persikabo 1973	34	4	8	22	44	74	-30	20

1.2 Menentukan Nilai Atribut

Pada tahap ini penelitian akan dilakukan dengan menggunakan *software Orange* dengan metode *Support Vector Machine (SVM)*. Data yang diolah menggunakan M.excel, sehingga di aplikasi *Orange* menggunakan operator *File* yang berfungsi untuk membaca data yang akan diolah. Kemudian masukan operator *Support Vector Machine* untuk membuat data yang diolah. Kemudian masukan operatot *predictions* untuk melihat hasil data yang diolah melalui *file* dan *Support Vector Machine*.



Gambar 4.1 Proses Awal *Software Orange*

1.3 Columns

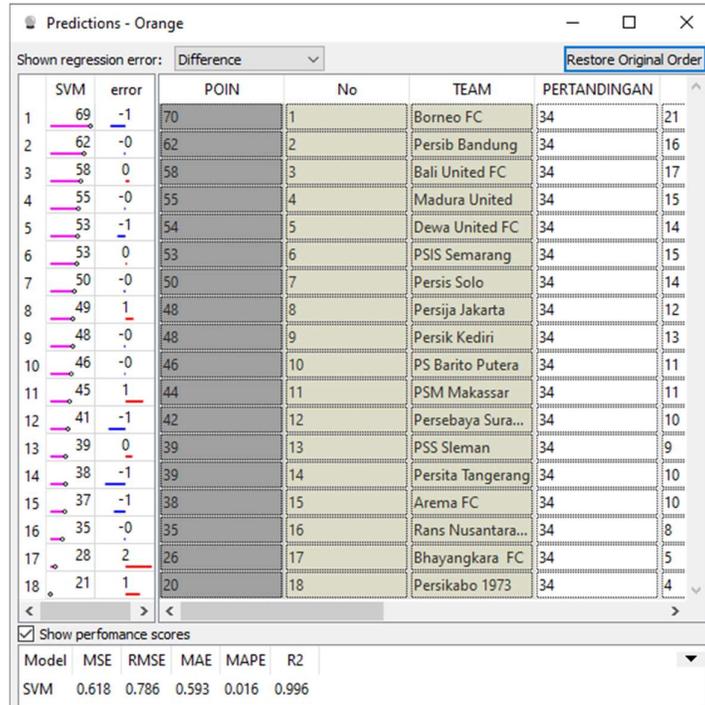
Columns disini berfungsi untuk menentukan target yang kita gunakan dalam menjalankan *software orange*. Target yang di tentukan dalam *columns* ialah “poin”.

Columns (Double click to edit)				
No	Name	Type	Role	Values
1	No	N numeric	meta	
2	PERTANDINGAN	N numeric	feature	
3	WIN	N numeric	feature	
4	DRAW	N numeric	feature	
5	LOS	N numeric	feature	
6	GF	N numeric	feature	
7	GA	N numeric	feature	
8	GD	N numeric	feature	
9	POIN	N numeric	target	
10	TEAM	S text	meta	

Gambar 4.2 Proses Columns

1.4 Predictions

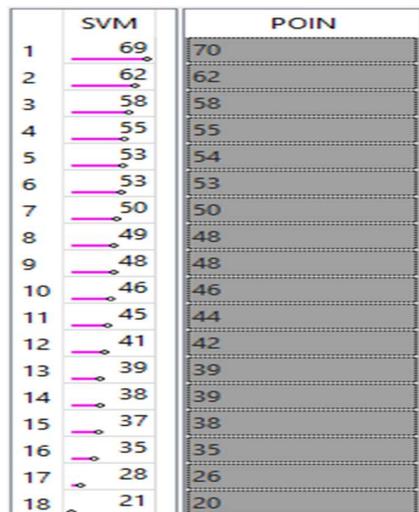
Predictions digunakan untuk memperlihatkan prediksi model hasil dari data yang di sambungkan dari File dan SVM. Disini didapatkan juga model SVM MSE sebesar 0,618, RMSE sebesar 0,786, MAE sebesar 0,593, MAPE sebesar 0,016, R2 sebesar 0,996.



Gambar 4.3 Hasil *Predictions*

1.5 Hasil

Disini menunjukkan hasil dari klasifikasi yang dapat dihasilkan kesimpulan. Berikut adalah hasil dari *Support Vector Machine (SVM)* menggunakan *software orange*.



Gambar 4.3 Hasil *Support Vector Machine*

KESIMPULAN DAN SARAN

1.6 Kesimpulan

Berdasarkan data yang dianalisis menggunakan *software* rakyat, Persebaya Surabaya tidak serta merta mendapatkan gelar juara karena klub tersebut memiliki nilai poin 42 dan nilai SVM 41. Alhasil, jika Persebaya Surabaya mencetak 69 poin atau lebih dalam satu musim, maka Persebaya Surabaya akan mendapatkan gelar juara. mereka mungkin akan memenangkan kejuaraan. Belum diketahui apa yang menjadi motivasi Persebaya Surabaya menjuarai BRI Liga 1, namun dipastikan tim tersebut tidak akan juara jika total poin akhirnya kurang dari 62. Persebaya Surabaya harus mencetak rata-rata lebih dari 54 gol per musim untuk bisa mencetak gol memenangkan BRI Liga jika mereka mencapai ini.

1.7 Saran

Untuk kedepannya data set yang akan diolah dengan *software* orange harus memiliki data yang lebih banyak agar hasil dari pengolahan data menggunakan *software* orange lebih optimal dan akurasi lebih tepat sehingga pengklasifikasi lebih sesuai.

DAFTAR PUSTAKA

- Bahri, S., & Lubis, A. (2020). Metode Klasifikasi Decision Tree Untuk Memprediksi Juara English Premier League. *Jurnal Sintaksis*, 2(1), 63-70.
- Hovi, H. S. W., Hadiana, A. I., & Umbara, F. R. (2022). Prediksi Penyakit Diabetes Menggunakan Algoritma Support Vector Machine (SVM). *Informatics and Digital Expert (INDEX)*, 4(1), 40-45.
- Nasrullah, A. H. (2021). Implementasi algoritma Decision Tree untuk klasifikasi produk laris. *Jurnal Ilmiah Ilmu Komputer Fakultas Ilmu Komputer Universitas Al Asyariah Mandar*, 7(2), 45-51.