

Prediksi Hasil Pertandingan Liga Serie A Menggunakan Metode Naïve Bayes

Ridwan Adi Pratama¹, Winda Apriandari², Didik Indrayana³

^{1,2,3} Teknik Informatika, Universitas Muhammadiyah Sukabumi, Kota Sukabumi, Indonesia

Email: ¹ridwanap24@ummi.ac.id, ²windaapriandari@ummi.ac.id, ³didikindrayana@ummi.ac.id

Email Penulis Korespondensi: ridwanap24@ummi.ac.id

Article History:

Received Jun 12th, 202x

Revised Aug 20th, 202x

Accepted Aug 26th, 202x

Abstrak

Sepak bola adalah olahraga yang sangat populer di seluruh dunia, dengan banyak penggemar yang melakukan prediksi hasil pertandingan. Metode yang digunakan untuk memprediksi hasil pertandingan dapat mempengaruhi akurasi prediksi tersebut. Dalam penelitian ini, dilakukan prediksi hasil pertandingan Liga Serie A menggunakan metode algoritma *Naïve Bayes*. Penelitian ini terdiri dari beberapa tahap, di antaranya adalah *Data Selection*, *Preprocessing Dataset*, *Transformation Dataset*, *Klasifikasi Naïve Bayes*, *Pengujian dan Evaluasi Model*. Data yang digunakan dalam penelitian ini diperoleh melalui *web scraping* dari situs www.football-data.uk, dan terdiri dari data Liga Serie A dari tahun 2017 hingga 2022. Setelah melalui tahap *preprocessing*, dataset diubah menjadi bentuk angka menggunakan metode *transformation dataset* agar dapat digunakan dalam algoritma *Naïve Bayes*. Kemudian dilakukan klasifikasi menggunakan metode *Naïve Bayes* untuk memprediksi hasil pertandingan. Hasil pengujian menunjukkan bahwa metode *Naïve Bayes* berhasil mencapai tingkat akurasi sebesar 75,79% dengan menggunakan data 1900 pertandingan. Empat teratas yaitu Juventus, Inter, Milan dan Napoli diprediksi akan lolos ke Liga *Champions*. Selanjutnya, model prediksi yang telah dibuat diimplementasikan dalam bentuk aplikasi. Aplikasi ini memungkinkan pengguna untuk melihat prediksi hasil pertandingan Liga Serie A dan menampilkan tampilan visual yang informatif. Dalam kesimpulan, penelitian ini berhasil melakukan prediksi hasil pertandingan Liga Serie A dengan menggunakan metode *Naïve Bayes*. Hasil pengujian menunjukkan tingkat akurasi yang baik. Penelitian ini memberikan kontribusi dalam pengembangan metode prediksi hasil pertandingan sepak bola dengan menggunakan algoritma *Naïve Bayes*.

Kata Kunci : Prediksi, Serie A, *Naïve Bayes*, Akurasi, *Web scrapping*

Abstract

Football is a highly popular sport worldwide, with many fans making predictions on match outcomes. The method used to predict match results can greatly impact the accuracy of the predictions. This research focuses on predicting the outcomes of matches in the Serie A league using the Naïve Bayes algorithm. The study consists of several stages, including Data Selection, Preprocessing Dataset, Transformation Dataset, Naïve Bayes Classification, and Model Testing and Evaluation. The data used in this research is obtained through web scraping from the website www.football-data.uk, comprising Serie A data from 2017 to 2022. After preprocessing, the dataset is transformed into numerical form using the transformation dataset method to be compatible with the Naïve Bayes algorithm. Classification is then performed using Naïve Bayes to predict match outcomes. The testing results indicate that the Naïve Bayes method achieves an accuracy rate of 75.79% using a dataset of 1900 matches. The top four teams, namely Juventus, Inter, Milan, and Napoli, are predicted to qualify for the Champions League. Furthermore, the developed prediction model is implemented as an application. This application allows users to view predictions of Serie A match outcomes and provides informative visual displays. In conclusion, this research successfully predicts the outcomes of Serie A matches using the Naïve Bayes method. The testing results demonstrate a good level of accuracy. This research

contributes to the development of football match prediction methods using the Naïve Bayes algorithm.

Keyword : Prediction, Serie A, Naïve Bayes, Accuracy, Web scrapping

1. PENDAHULUAN

Sepak bola merupakan olahraga yang sangat populer di seluruh dunia karena sekitar 4% dari total populasi dunia terlibat secara aktif dalam permainan tersebut [1]. Sepak bola adalah olahraga yang sangat banyak penganutnya di dunia, Sepak bola bahkan dapat berdampak bagi kesehatan, kegiatan mata pencaharian maupun bisnis [2]. Sepak bola merupakan cabang olahraga yang dimainkan oleh 2 tim, setiap tim terdiri 11 pemain dan bertanding kandang dengan tandang dengan durasi bermain 2x45 menit [3]. Salah satu Liga yang paling banyak digemari yaitu kompetisi Liga Champions.

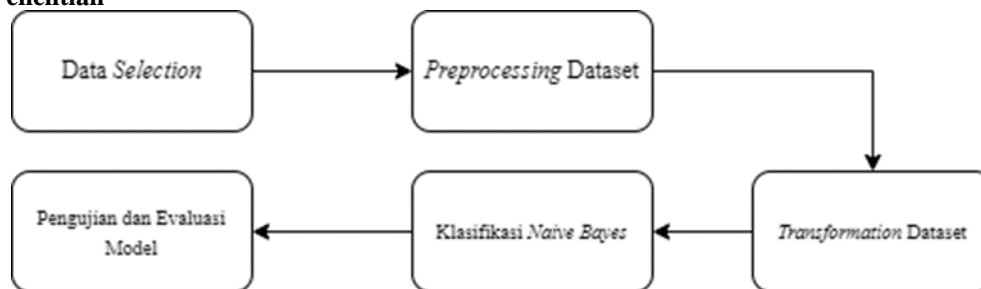
Liga Champions merupakan kompetisi sepakbola yang diselenggarakan setiap tahun oleh Union of European Football Associations (UEFA) dan yang mengikuti kompetisi ini adalah 4 teratas pada masing-masing Liga di Eropa [4]. Salah satu Liga yang mengikuti Liga Champions yaitu Serie A. Serie A atau yang disebut juga Liga Italia merupakan liga sepakbola paling tertinggi di Italia yang mempunyai 20 tim sepakbola terbaik di Italia. Dengan 20 tim, 4 peringkat teratas akan bertanding untuk memperebutkan Liga Champions [5].

Banyak sekali penganut melakukan prediksi sepak bola terutama pecinta tim-tim kesayangannya. Prediksi adalah suatu proses untuk memproyeksikan atau mengestimasi nilai variabel di masa depan. Prediksi tidak selalu menghasilkan jawaban yang pasti tentang kejadian yang akan terjadi, tetapi berusaha untuk mendekati sebanyak mungkin dengan apa yang terjadi [6]. Terlihat banyak sekali penganut melakukan prediksi tetapi karena keegoisan sendiri tidak memikirkan variabel apa yang terjadi di masa lalu. Dengan begitu prediksi hasil pertandingan dibutuhkan sebuah metode agar dalam hal memprediksi menjadi lebih baik. Metode yang digunakan untuk memprediksi hasil pertandingan menggunakan variabel tertentu yaitu dengan metode algoritma *naïve bayes*.

Naïve bayes adalah salah satu metode yang dapat digunakan untuk melakukan klasifikasi data. Klasifikasi *naïve bayes* adalah teknik statistik yang dapat digunakan untuk memperkirakan probabilitas pada suatu kelas [7]. Probabilitas adalah suatu estimasi mengenai kemungkinan yang akan terjadi pada suatu peristiwa dengan melakukan prediksi terhadap apa yang akan terjadi. Rentang probabilitas berkisar antara 0 dan 1. Jika probabilitas dengan nilai 0, maka peristiwa tersebut tidak akan terjadi, sedangkan jika probabilitas dengan nilai 1 maka peristiwa tersebut diprediksi akan terjadi [8]. Dengan menggunakan metode *naïve bayes* yang menghitung dengan probabilitas yang melibatkan variabel-variabel yang tepat untuk memprediksi hasil pertandingan liga Serie A.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian



Gambar 1. Tahapan Penelitian

- Data Selection**
Data selection adalah proses mengurangi jumlah data yang digunakan dalam proses *mining*, dengan tetap mempertahankan representasi data asli [9]. Data yang digunakan penulis yaitu Liga Serie A dari tahun 2017-2022 dengan menggunakan *web scraping*. *Web scraping* adalah proses pengambilan informasi atau dokumen semi-terstruktur dari internet [10]. Dataset diperoleh dari www.football-data.uk.
- Preprocessing dataset**
Preprocessing merupakan tahap dimana variabel yang diterapkan dihilangkan yang tidak relevan dalam proses pengolahan selanjutnya [11]. Variabel atau atribut yang digunakan yaitu HT, AT, HTHG, HTAG, HS, AS, HST, AST, HF, AF, HC, AC, HY, AY, HR, AR dan FTR.
- Transformation dataset**
Transformation dataset adalah proses mengubah nilai-nilai dalam dataset yang awalnya dalam bentuk huruf menjadi bentuk angka, karena *naïve bayes* lebih cepat menerima data dalam bentuk angka [12].

d. Klasifikasi *Naïve bayes*

Klasifikasi *naïve bayes* adalah salah satu metode yang digunakan untuk pengambilan keputusan dalam masalah klasifikasi, yang dapat menghasilkan hasil yang terbaik [13]. Teorema bayes adalah suatu teorema dalam bidang statistika yang digunakan untuk menghitung probabilitas suatu hipotesis [14]. Rumus Probabilitas yaitu :

$$P(H|X) = \frac{P(X|H) \times P(H)}{P(X)} \tag{1}$$

Keterangan :

H : Sampel data *training*.

X : Atributur untuk data *training*.

P(H|X) : Probabilitas H berdasarkan atribut X (*posterior probability*).

P(X|H) : Probabilitas atribut X berdasarkan H (*likelihood*).

P(H) : Probabilitas H (*prior probability*).

P(X) : Probabilitas atribut X.

e. Pengujian dan Evaluasi Model

Proses pengujian yang dilakukan yaitu menggunakan data *training* dan data *testing* dengan persentase (80% : 20%). Dengan menganalisis kinerja yang digunakan yaitu akurasi, *recall*, presisi dan *f-score*. Evaluasi model yaitu melakukan *re-sampling* dengan melihat seberapa baik model dapat tergeneralisasikan untuk mencegah *overfitting* dan *underfitting*. *Overfitting* adalah situasi di mana model yang telah melalui proses pelatihan mampu mencapai tingkat akurasi yang tinggi pada sebagian besar data latihan, tetapi mengalami kesalahan yang signifikan saat digunakan untuk melakukan prediksi [15]. Sedangkan *underfitting* adalah situasi kondisi di mana model tidak mampu menangkap logika atau pola yang terdapat dalam data, sehingga sulit untuk melakukan prediksi dengan akurat baik pada dataset pelatihan maupun dataset pengujian [16]. Dengan demikian pada penelitian ini menggunakan teknik *K-fold Cross Validation*. *K-Fold Cross Validation* digunakan sebagai alat untuk memperkirakan kesalahan yang ada pada prediksi model serta mengevaluasi performanya [17].

2.2 Machine Learning

Machine Learning adalah sebuah implementasi dari kecerdasan buatan (*Artificial Intelligence*) yang berfokus pada pengembangan sistem yang memiliki kemampuan untuk belajar sendiri tanpa diprogram secara berulang. Dengan kata lain, *machine learning* yaitu pemrograman komputer yang bertujuan mencapai performa tertentu dengan menggunakan kumpulan data *training* atau pengalaman di masa lalu [18].

3. HASIL DAN PEMBAHASAN

3.1 Data Selection

Data yang digunakan penulis yaitu Liga *Serie A* dari tahun 2017-2022 dengan menggunakan *web scraping* yang diperoleh dari www.football-data.uk.

3.2 Preprocessing Dataset

Dataset yang telah didapatkan dari data *selection* dipilih variabel yang sesuai dengan penelitian dan menghilangkan data yang tidak terpakai. Variabel atau atribut yang digunakan yaitu HT, AT, HTHG, HTAG, HS, AS, HST, AST, HF, AF, HC, AC, HY, AY, HR, AR dan FTR. Seperti pada gambar berikut :

	HomeTeam	AwayTeam	FTHG	FTAG	FTR	HTHG	HTAG	HTR	HS	AS	HST	AST	\
0	Juventus	Cagliari	3	0	H	2	0	H	17	5	7	2	
1	Verona	Napoli	1	3	A	0	2	A	12	22	4	12	
2	Atalanta	Roma	0	1	A	0	1	A	14	6	6	1	
3	Bologna	Torino	1	1	D	1	1	D	13	9	3	3	
4	Crotone	Milan	0	3	A	0	3	A	6	19	2	10	
...
375	Milan	Fiorentina	5	1	H	2	1	H	18	9	11	3	
376	Napoli	Crotone	2	1	H	2	0	H	19	9	9	4	
377	Sassuolo	Roma	0	1	A	0	1	A	16	15	3	6	
378	Spal	Sampdoria	3	1	H	1	0	H	16	9	6	3	
379	Udinese	Bologna	1	0	H	1	0	H	20	8	5	5	
	HF	AF	HC	AC	HY	AY	HR	AR					
0	13	11	4	1	1	0	0	0					
1	7	9	3	6	2	2	0	1					
2	14	16	4	1	1	2	0	0					
3	18	10	0	8	2	1	0	0					
4	9	14	2	9	1	3	1	0					
...					
375	10	5	6	1	2	1	0	1					
376	3	9	6	9	0	0	0	0					
377	14	8	5	9	1	1	0	0					
378	10	11	5	3	1	4	0	0					
379	9	11	10	2	0	2	0	0					

Gambar 2. Dataset yang digunakan

3.3 Transformation Dataset

Transformation dataset ini mengubah nilai dataset yang awalnya berupa huruf menjadi angka. Atribut yang diubah yaitu *HomeTeam* dan *AwayTeam* serta pada Label FTR. Berikut tabel yang telah diubah :

Tabel 1. *HomeTeam* dan *AwayTeam* dalam bentuk numerik

Atribut	Hasil Numerik
<i>Roma</i>	1
<i>Inter</i>	2
<i>Benevento</i>	3
<i>Spal</i>	4
<i>Genoa</i>	5
<i>Crotone</i>	6
<i>Napoli</i>	7
<i>Fiorentina</i>	8
<i>Chievo</i>	9
<i>Sampdoria</i>	10
<i>Bologna</i>	11
<i>Cagliari</i>	12
<i>Udinese</i>	13
<i>Sassuolo</i>	14
<i>Verona</i>	15
<i>Torino</i>	16
<i>Lazio</i>	17
<i>Atalanta</i>	18
<i>Parma</i>	19
<i>Lecce</i>	20
<i>Empoli</i>	21
<i>Juventus</i>	22
<i>Milan</i>	23
<i>Frosinone</i>	24
<i>Brescia</i>	25
<i>Spezia</i>	26
<i>Salernitana</i>	27
<i>Venezia</i>	28

Tabel 2. Label FTR

Label	Hasil Numerik
Home	1
Away	2
Draw	0

3.4 Klasifikasi Naïve Bayes dan Pengujian

a. *Encoding data*

Pada *HomeTeam* dan *AwayTeam* sudah diubah ke dalam bentuk numbering atau *integer* yang menghasilkan sebagai berikut.

```
# Import dataset
dataset = pd.read_csv('Data_2017-2022.csv', sep=";")
print(dataset)

HomeTeam  AwayTeam  FTHG  FTAG  HTHG  HTAG  HS  AS  HST  AST  HF  AF  \
0         1         12     3     0     2     0  17  5   7   2  13  11
1         15         5     1     3     0     2  12  22  4  12  7   9
2         6         4     0     1     0     1  14  6   6   1  14  16
3         11        16     1     1     1     1  13  9   3   3  18  10
4         18         3     0     3     0     3   6  19   2  10  9  14
...
1895        26         5     0     3     0     3  16  10  4   4  10  9
1896         2        10     3     0     0     0  27  9  12   3   6  12
1897        14         3     0     3     0     3  15  18   3  12  8  13
1898        27        13     0     4     0     3  25  19  10  11  11  12
1899        28        12     0     0     0     0   7  31   1   7   6  19

HC  AC  HY  AY  HR  AR  FTR
0   4   1   1   0   0   1
1   3   6   2   2   0   2
2   4   1   1   2   0   2
3   0   8   2   1   0   0
4   2   9   1   3   1   2
...
1895  4   2   1   2   0   2
1896  8   2   1   2   0   1
1897  3   7   2   1   0   2
1898  7   1   1   1   0   2
1899  6  11   0   3   0   0

[1900 rows x 19 columns]
```

Gambar 3. Dataset dalam bentuk *integer*

b. *Pembagian data training dan data testing*

Dataset dibagi menjadi 2 yaitu data *training* dan data *testing* dengan pembagian dataset yaitu data *training* 80% dan data *testing* 20% yang dilakukan pembagian data menggunakan *train test split* pada gambar berikut.

```
# Splitting the dataset into the Training set and Test set
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.20, random_state=0)
```

Gambar 4. Pembagian dataset menggunakan *train test split*

c. *Prediksi Naïve Bayes*

Hasil yang dilakukan menggunakan algoritma *naïve bayes* yaitu mendapatkan hasil sebesar 75,79% dengan menggunakan data 1900 pertandingan dari Liga *Seire A* pada tahun 2017/2018 hingga 2021/2022. Persentase akurasi dihasilkan dengan menghitung data latih penelitian pada gambar berikut.

```
# Prediksi Test set results
y_pred = classifier.predict(X_test)
print(y_pred)

[2 2 2 0 1 2 2 1 1 0 2 1 2 0 1 1 1 2 1 2 1 1 0 2 1 0 2 0 1 1 1 1 0 2 2 2 0
 1 0 1 0 0 1 1 2 2 1 0 2 0 0 1 0 0 2 1 1 2 0 1 0 0 1 0 2 0 1 1 2 2 2 2 1 2
 2 1 1 0 1 1 1 2 2 1 1 2 0 2 2 0 1 1 1 0 1 1 1 2 0 1 0 1 1 0 0 0 2 2 2 1 1
 0 1 0 2 1 2 1 2 2 1 0 2 2 2 2 0 1 0 2 1 0 1 0 0 0 1 1 0 0 2 1 1 0 1 2 2 0
 1 1 1 0 1 2 1 0 2 2 1 2 1 1 0 2 2 2 2 0 0 1 1 0 1 2 2 0 1 1 1 1 0 2 0 2 0
 1 2 2 2 2 1 0 1 1 2 2 2 2 2 1 0 1 0 2 0 2 2 1 1 2 1 1 0 0 1 1 2 2 0 0
 2 1 2 0 0 1 1 1 2 2 0 1 0 1 1 0 0 0 1 1 1 1 1 0 0 0 2 1 2 2 2 0 2 1 0 1 0
 2 1 2 1 1 2 2 1 2 2 0 2 2 1 0 0 0 1 1 1 1 0 0 2 1 2 0 0 1 2 0 2 2 1 0 1 2
 1 2 0 1 2 0 2 0 1 2 0 2 1 1 1 0 2 0 0 0 0 1 2 1 1 0 2 0 1 1 2 2 1 2 0 0 0
 1 1 1 0 0 1 2 2 1 0 1 1 2 0 1 0 1 1 2 0 0 1 1 1 0 0 2 1 0 1 2 2 0 1 1 1
 1 2 1 2 1 2 0 0 0 2]
```

Gambar 5. Prediksi data *testing*

```
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, y_pred)
print(cm)

[[ 65  19  21]
 [ 29 124   2]
 [ 18   3  99]]

from sklearn.metrics import classification_report
print(classification_report(y_test, y_pred))

              precision    recall  f1-score   support

   0             0.58       0.62       0.60         105
   1             0.85       0.80       0.82         155
   2             0.81       0.82       0.82         120

 accuracy                   0.76         380
 macro avg                   0.75         380
weighted avg                   0.76         380

from sklearn.metrics import accuracy_score
print('Accuracy: ', accuracy_score(y_test, y_pred)*100)

Accuracy: 75.78947368421053

from sklearn.model_selection import cross_val_score
accuracies = cross_val_score(estimator = classifier, X = X_train, y = y_train, cv = 10)
print (accuracies.mean()*100)

71.77631578947367
```

Gambar 6. Akurasi yang dihasilkan

Pada pengujian ini menghitung akurasi, presisi, *recall* dan *f1-score* yang terdapat pada gambar diatas. Dengan menggunakan *cross validation* dengan akurasi 71,78%.

Selanjutnya menghitung klasemen akhir dengan membuat kalkulasi *point*. *Home win* akan mendapatkan 3 *point*, *Away win* akan mendapatkan 3 *point* dan *Draw* akan mendapatkan 1 *point*. Berikut kalkulasi serta *final standing* terdapat pada gambar dibawah.

```
# Kalkulasi final point by team
points = {}
for i in range(len(result)):
    home = result[i][0]
    away = result[i][1]
    pred = result[i][2]
    if home not in points:
        points[home] = 0
    if away not in points:
        points[away] = 0
    if pred == 0:
        points[home] += 0
        points[away] += 3
    elif pred == 1:
        points[home] += 1
        points[away] += 1
    else:
        points[home] += 3
        points[away] += 0
```

Gambar 7. Kalkulasi Point

```
#Sorting teams by points
sorted_points = sorted(points.items(), key=lambda x: x[1], reverse=True)
```

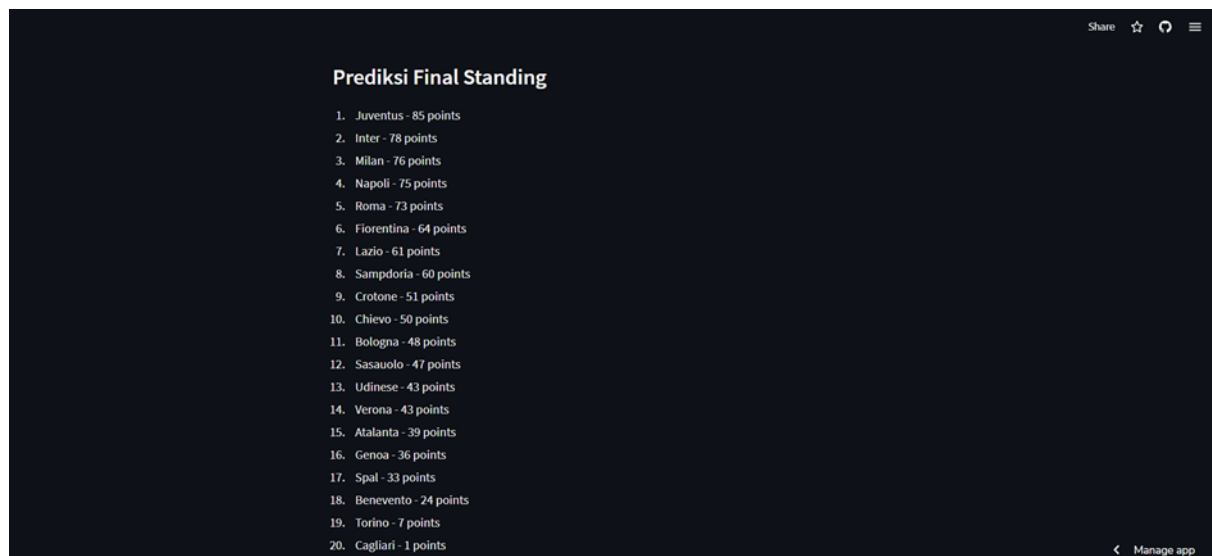
Gambar 8. Sorting team by point

```
#Print final standings
print("Final Standings:")
print("Team\tPoints")
for i in range(len(sorted_points)):
    print(f"{sorted_points[i][0]}\t{sorted_points[i][1]}")
```

```
Final Standings:
Team    Points
Juventus    85
Inter      78
Milan      76
Napoli     75
Roma       73
Fiorentina  64
Lazio      61
Sampdoria  60
Crotone    51
Chievo     50
Bologna    48
Sasuolo    47
Udinese    43
Verona     43
Atalanta  39
Genoa      36
Spal       33
Benevento  24
Torino     7
Cagliari   1
```

Gambar 9. Final Standing

Dari hasil pengujian tersebut diimplementasikan ke dalam bentuk aplikasi yang membuat semua orang dapat melihat prediksi hasil pertandingan liga *Serie A*. Berikut tampilan pada aplikasi tersebut.



Prediksi Final Standing

1. Juventus - 85 points
2. Inter - 78 points
3. Milan - 76 points
4. Napoli - 75 points
5. Roma - 73 points
6. Fiorentina - 64 points
7. Lazio - 61 points
8. Sampdoria - 60 points
9. Crotone - 51 points
10. Chievo - 50 points
11. Bologna - 48 points
12. Sasuolo - 47 points
13. Udinese - 43 points
14. Verona - 43 points
15. Atalanta - 39 points
16. Genoa - 36 points
17. Spal - 33 points
18. Benevento - 24 points
19. Torino - 7 points
20. Cagliari - 1 points

Gambar 10. Tampilan Prediksi *Final Standing*



Prediksi Histori Match Result

1 = Home Win
2 = Away Win
0 = Draw

	Home Team	Away Team	Result
0	Napoli	Crotone	2
1	Atalanta	Milan	2
2	Milan	Spal	2
3	Crotone	Inter	0
4	Lazio	Roma	1
5	Bologna	Udinese	2
6	Chievo	Juventus	2
7	Milan	Juventus	1
8	Juventus	Atalanta	1
9	Spal	Verona	0

Gambar 11. Tampilan Prediksi Hasil Pertandingan

4. KESIMPULAN

Pada penelitian ini melakukan prediksi hasil pertandingan dengan klasemen akhir pada liga *Seire A* dengan menggunakan metode *naïve bayes* mendapatkan akurasi sebesar 75,79%. Empat tim teratas yaitu Juventus, Inter, Milan dan Napoli diprediksi akan lolos ke Liga *Champions*. Data yang diambil menggunakan *web scraping*, dataset tersebut yaitu liga *Serie A* pada tahun 2017/2018 hingga 2021/2022. Dengan hasil pengujian ini *naïve bayes* mampu melakukan persentase akurasi yang lebih baik dibandingkan dengan metode lainnya yang sudah diuji oleh pengujian sebelumnya.

UCAPAN TERIMA KASIH

Terima kasih kepada Allah SWT yang telah memberikan kemudahan, kesabaran serta ketabahan terhadap penulis, selanjutnya kepada orang tua atas do'a dan dukungannya yang membuat penulis menyelesaikan penelitian ini, juga kepada dosen pembimbing saya yang selalu memberikan saran dan arahan terhadap penelitian ini sehingga menjadi sebuah penelitian yang terarah dan tepa tapa yang penulis inginkan, dan yang terakhir terima kasih kepada teman-teman saya yang selalu memberikan ilmu tambahan hingga memberikan ilmu baru yang belum pernah mengerti oleh penulis yang berdampak yaitu wawasan penulis menjadi luas.

DAFTAR PUSTAKA

- [1] S. N. Fajrin, A. Agustiyawan, P. Purnamadyawati, and D. S. Mahayati, "Literature Review : Hubungan Koordinasi Terhadap Keterampilan Menggiring Bola Pada Pemain Sepak Bola," *Indones. J. Physiother.*, vol. 1, no. 1, p. 6, 2021, doi: 10.52019/ijpt.v1i1.2605.
- [2] A. S. Yudistira and A. Nugroho, "PREDICTION OF THE ENGLISH PREMIER LEAGUE CHAMPION

- TEAM FOR THE 2021 / 2022 SEASON USING THE NAÏVE BAYES METHOD MENGGUNAKAN METODE NAÏVE BAYES,” vol. 3, no. 5, pp. 1239–1243, 2022.
- [3] N. Priyo Utomo and P. Indarto, “Analisis Keterampilan Teknik Dasar Passing dalam Sepak Bola,” *J. Porkes*, vol. 4, no. 2, pp. 87–94, 2021, doi: 10.29408/porkes.v4i2.4578.
- [4] C. Tandian, Y. Laia, and A. Saputra, “Penerapan Data Mining Dalam Memprediksi Pemenang Klub Sepak Bola Pada Ajang Liga Champion Dengan Algoritma C.45,” *J. Sist. Inf. dan Ilmu Komput. Prima(JUSIKOM PRIMA)*, vol. 2, no. 2, pp. 91–98, 2019, doi: 10.34012/jusikom.v2i2.397.
- [5] N. N. Fadrijin and A. Wibawa, “Pemodelan Deret Waktu Point Liga Italia Serie a Dengan Pendekatan Regresi Berdasarkan Rmse (Root Mean Square Score) Terkecil Dan Skor Maksimal Tiap Pekan,” *J. Stat. Univ. Muhammadiyah Semarang*, vol. 8, no. 1, pp. 78–87, 2020, [Online]. Available: <https://jurnal.unimus.ac.id/index.php/statistik/article/view/5802>
- [6] B. Putro, M. T. Furqon, and S. H. Wijoyo, “Prediksi Jumlah Kebutuhan Pemakaian Air Menggunakan Metode Exponential Smoothing (Studi Kasus : PDAM Kota Malang),” vol. 2, no. 11, pp. 4679–4686, 2018.
- [7] J. Teknologi, I. Jtsi, M. R. Handoko, F. Teknik, and U. T. Indonesia, “SISTEM PAKAR DIAGNOSA PENYAKIT SELAMA KEHAMILAN MENGGUNAKAN METODE NAIVE BAYES BERBASIS WEB,” vol. 2, no. 1, pp. 50–58, 2021.
- [8] H. Abijono, P. Santoso, and N. L. Anggreini, “Algoritma Supervised Learning Dan Unsupervised Learning Dalam Pengolahan Data,” *J. Teknol. Terap. G-Tech*, vol. 4, no. 2, pp. 315–318, 2021, doi: 10.33379/gtech.v4i2.635.
- [9] F. A. Pratama, R. Narasati, and D. R. Amalia, “Pengaruh Kata Cashback Terhadap Peningkatan Penjualan Menggunakan Data Mining,” *J. Ilm. Manaj. Inform. dan Komput.*, vol. 3, no. 2, pp. 1–5, 2019.
- [10] D. F. Setiawan, T. Tristiyanto, and A. Hijriani, “Aplikasi Web Scraping Deskripsi Produk,” *J. Teknoinfo*, vol. 14, no. 1, p. 41, 2020, doi: 10.33365/jti.v14i1.498.
- [11] H. Najjichah, A. Syukur, and H. Subagyo, “Pengaruh Text Preprocessing Dan Kombinasinya Pada Peringkat Dokumen Otomatis Teks Berbahasa Indonesia,” *J. Teknol. Inf.*, vol. XV, no. 1, pp. 1–11, 2019.
- [12] R. Hayami, Soni, and I. Gunawan, “Klasifikasi Jamur Menggunakan Algoritma Naïve Bayes,” *J. CoSciTech (Computer Sci. Inf. Technol.)*, vol. 3, no. 1, pp. 28–33, 2022, doi: 10.37859/coscitech.v3i1.3685.
- [13] D. Alita, I. Sari, and A. Rahman Isnain, “Penerapan Naïve Bayes Classifier Untuk Pendukung Keputusan Penerima Beasiswa,” *Jdmsi*, vol. 2, no. 1, p. 702022, 2021.
- [14] N. Sulardi and A. Witanti, “Sistem Pakar Untuk Diagnosis Penyakit Anemia Menggunakan Teorema Bayes,” *J. Tek. Inform.*, vol. 1, no. 1, pp. 19–24, 2020, doi: 10.20884/1.jutif.2020.1.1.12.
- [15] P. A. Nugroho, I. Fenriana, and R. Arijanto, “Implementasi Deep Learning Menggunakan Convolutional Neural Network (Cnn) Pada Ekspresi Manusia,” *Algor*, vol. 2, no. 1, pp. 12–21, 2020.
- [16] S. Muhammad and A. T. Wibowo, “Klasifikasi Tanaman Aglaonema Berdasarkan Citra Daun Menggunakan Metode Convolutional Neural Network (Cnn),” *e-Proceeding Eng.*, vol. 8, no. 5, pp. 10621–10636, 2021.
- [17] L. Mardiana, D. Kusnandar, and N. Satyahadewi, “Analisis Diskriminan Dengan K Fold Cross Validation Untuk Klasifikasi Kualitas Air Di Kota Pontianak,” *Bimaster Bul. Ilm. Mat. Stat. dan Ter.*, vol. 11, no. 1, pp. 97–102, 2022, [Online]. Available: <https://jurnal.untan.ac.id/index.php/jbmstr/article/view/51608>
- [18] C. Chazar and B. Erawan, “Machine Learning Diagnosis Kanker Payudara Menggunakan Algoritma Support Vector Machine,” *Inf. (Jurnal Inform. dan Sist. Informasi)*, vol. 12, no. 1, pp. 67–80, 2020, doi: 10.37424/informasi.v12i1.48.