

**PREDIKSI POSISI TIM NASIONAL INDONESIA DI AKHIR
ROUND 3 KUALIFIKASI PIALA DUNIA GRUP C
MENGUNAKAN *MACHINE LEARNING***

SKRIPSI

DISUSUN OLEH

TAMAM AULIYA HABIB

NPM. 2109020154



UMSU

Unggul | Cerdas | Terpercaya

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
MEDAN
2025**

**PREDIKSI POSISI TIM NASIONAL INDONESIA DI AKHIR ROUND 3
KUALIFIKASI PIALA DUNIA GRUP MENGGUNAKAN *MACHINE
LEARNING***

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer
(S.Kom) dalam Program Studi Teknologi Informasi pada Fakultas Ilmu Komputer
dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara**

TAMAM AULIYA HABIB

NPM. 2109020154

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
MEDAN
2025**

LEMBAR PENGESAHAN

Judul Skripsi : PREDIKSI POSISI TIM NASIONAL INDONESIA DI
AKHIR ROUND 3 KUALIFIKASI PIALA DUNIA
GRUP C MENGGUNAKAN *MACHINE LEARNING*
Nama Mahasiswa : Tamam Auliya Habib
NPM : 2109020154
Program Studi : Teknologi Informasi

Menyetujui
Komisi Pembimbing



Halim Maulana, S.T., M.Kom
NIDN. 0121119102

Ketua Program Studi



Fatma Sari Hutagalung, S.Kom., M.Kom
NIDN. 0117019301

Dekan



Dr. Al Kibwarizmi, S.Kom., M.Kom.
NIDN. 0127099201

PERNYATAAN ORISINALITAS

PREDIKSI POSISI TIM NASIONAL INDONESIA DI AKHIR ROUND 3 KUALIFIKASI PIALA DUNIA GRUP C MENGGUNAKAN *MACHINE LEARNING*

SKRIPSI

Saya menyatakan bahwa karya tulis ini adalah hasil karya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing disebutkan sumbernya.

Medan, 30 Desember 2024

Yang membuat pernyataan



Tamam Auliya Habib

NPM. 2109020154

PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika Universitas Muhammadiyah Sumatera Utara, saya bertanda tangan dibawah ini:

Nama : Tamam Auliya Habib
NPM : 2109020154
Program Studi : Teknologi Informasi
Karya Ilmiah : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Muhammadiyah Sumatera Utara Hak Bedas Royalti Non-Eksekutif (*Non-Exclusive Royalty free Right*) atas penelitian skripsi saya yang berjudul:

Prediksi Posisi Tim Nasional Indonesia di Akhir Round 3 Kualifikasi Piala Dunia Grup Menggunakan *Machine learning*

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-Eksekutif ini, Universitas Muhammadiyah Sumatera Utara berhak menyimpan, mengalih media, memformat, mengelola dalam bentuk database, merawat dan mempublikasikan Skripsi saya ini tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemegang dan atau sebagai pemilik hak cipta.

Demikian pernyataan ini dibuat dengan sebenarnya.

Medan, 30 Desember 2024

Yang membuat pernyataan



Tamam Auliya Habib

NPM. 2109020154

RIWAYAT HIDUP

DATA PRIBADI

Nama Lengkap : Tamam Auliya Habib
Tempat dan Tanggal Lahir : Medan, 09 Juli 2003
Alamat Rumah : Jln.Musholla Gg.Sukma No.2
Telepon/Faks/HP : 0895618114639
E-mail : tamamumsu@gmail.com
Instansi Tempat Kerja : -
Alamat Kantor : -

DATA PENDIDIKAN

SD : Free Methodist 2 Medan TAMAT: 2016
SMP : SMP Panca Budi Medan TAMAT: 2019
SMA : SMA Panca Budi Medan TAMAT: 2021

KATA PENGANTAR



Puji syukur Penulis ucapkan kepada Allah SWT atas rahmat-Nya Penulis dapat menyelesaikan skripsi ini. Penulisan skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer pada Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara.

Penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada kedua orang tua yang telah memberikan dukungan moril maupun materil kepada Penulis. Dan Penulis menyadari bahwa tanpa bantuan, bimbingan, dan dukungan dari berbagai pihak sejak masa perkuliahan hingga masa penyusunan skripsi, sulit untuk menyelesaikan skripsi ini. Oleh karena itu dengan segala kerendahan hati, Penulis ingin menyampaikan terima kasih dan rasa hormat kepada:

1. Bapak Prof. Dr. Agussani, M.AP., Rektor Universitas Muhammadiyah Sumatera Utara (UMSU)
2. Bapak Dr. Al-Khowarizmi, S.Kom., M.Kom. Dekan Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
3. Bapak Halim Maulana, S.T., M.Kom, selaku Wakil Dekan I Fakultas Ilmu Komputer dan Teknologi Informasi dan juga selaku Dosen Pembimbing Penulis.
4. Bapak Dr. Lutfi Basit, S.Sos., M.I.Kom, selaku Wakil Dekan III Fakultas Ilmu Komputer dan Teknologi Informasi.

5. Ibu Fatma Sari Hutagalung, S.Kom., M.Kom, selaku Ketua Program Studi Teknologi Informasi yang telah memberikan bimbingan dan bantuan selama Penulis menempuh studi serta yang selalu mengingatkan Penulis untuk segera menyelesaikan skripsi ini.
6. Bapak Mhd. Basri, S.Si., M.Kom, selaku Sekretaris Program Studi Teknologi Informasi
7. Ibu Ibu Fatma Sari Hutagalung, S.Kom., M.Kom, selaku Dosen Penguji I dan Bapak Dr. Zainal Aziz. M.Si selaku Dosen Penguji II yang telah membantu dan memberikan saran dalam pengerjaan skripsi ini.
8. Kedua orang tua saya yang selalu mensupport dengan cara apapun dan dengan keadaan apapun
9. Kepada sahabat saya, Kinoy, yang telah banyak membantu selama proses penyusunan tugas ini. Terima kasih atas segala waktu, tenaga, dan pengorbanan yang diberikan, mulai dari bolak-balik ke kampus, menalangi biaya cetak, hingga mendampingi saya dalam berbagai situasi sulit.
10. Semua pihak yang terlibat langsung ataupun tidak langsung yang tidak dapat penulis ucapkan satu-persatu yang telah membantu penyelesaian skripsi ini.

Medan, 30 Desember 2024

Penulis

Tamam Auliya Habib

**PREDIKSI POSISI TIM NASIONAL INDONESIA DI AKHIR ROUND 3
KUALIFIKASI PIALA DUNIA GRUP C TIM MENGGUNAKAN
*MACHINE LEARNING***

ABSTRAK

Sepak bola memiliki popularitas tinggi di Indonesia, namun Tim Nasional menghadapi tantangan besar dalam Kualifikasi Piala Dunia 2026, terutama di Putaran 3 Grup C, akibat kesenjangan peringkat FIFA dengan tim lain. Penelitian ini bertujuan untuk memprediksi posisi akhir Tim Nasional Indonesia dalam grup tersebut menggunakan metode machine learning, khususnya Support Vector Machine (SVM), berdasarkan data statistik pertandingan. Proses penelitian mencakup pengumpulan data, preprocessing, pelatihan model, dan evaluasi akurasi prediksi. Hasil menunjukkan bahwa model memiliki tingkat akurasi 67%, dengan probabilitas Indonesia menempati posisi 2 sebesar 19.73% dan menetap di posisi 4 sebesar 17.43%. Hasil ini memberikan gambaran realistis mengenai peluang Indonesia dalam persaingan grup dan menunjukkan bagaimana pendekatan berbasis data dapat digunakan untuk memproyeksikan kemungkinan hasil dalam kompetisi sepak bola.

Kata Kunci : sepak bola, *Machine Learning*, *Support Vector Machine*, prediksi, kualifikasi piala dunia

PREDICTION OF INDONESIAN NATIONAL TEAM POSITION AT THE END OF ROUND 3 OF WORLD CUP QUALIFICATION GROUP C USING *MACHINE LEARNING*

ABSTRACT

Football has a high popularity in Indonesia, but the National Team faces a big challenge in the 2026 World Cup Qualifiers, especially in Round 3 of Group C, due to the FIFA ranking against other teams. This study aims to predict the final position of the Indonesian National Team in the group using machine learning methods, specifically Support Vector Machine (SVM), based on match statistics data. The research process includes data collection, preprocessing, model training, and evaluation of prediction accuracy. The results show that the model has an accuracy rate of 67%, with the probability of Indonesia occupying position 2 of 19.73% and settling in position 4 of 17.43%. These results provide a realistic picture of Indonesia's chances in the group competition and show how a data-driven approach can be used to project possible outcomes in a football competition.

Keywords: Football, *Machine Learning*, *Support Vector Machine*, Predictions, World Cup Qualifiers

DAFTAR ISI

KATA PENGANTAR	iv
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
BAB II LANDASAN TEORI	5
2.1 Sepak Bola.....	5
2.2 <i>Machine Learning</i>	6
2.3 <i>SVM (Support Vector Machine)</i>	7
2.4 <i>Python</i>	8
2.5 <i>Classification Report</i>	8
2.6 Penelitian Terkait.....	9
BAB III METODOLOGI PENELITIAN	15
3.1 Pendekatan Penelitian	15
3.2 Teknik Pengumpulan Data	15
3.3 Tahapan Penelitian.....	17
3.4 Alur Pengolahan Data	19
BAB IV HASIL DAN PEMBAHASAN	21
BAB V PENUTUP	30
5.1 Kesimpulan.....	30
5.2 Saran	30
DAFTAR PUSTAKA	32

DAFTAR TABEL

Tabel 2.1	13
-----------------	----

DAFTAR GAMBAR

Gambar 3.1	17
Gambar 3.2	19
Gambar 4.1	21
Gambar 4.2	22
Gambar 4.3	22
Gambar 4.4	23
Gambar 4.5	23
Gambar 4.6	24
Gambar 4.7	24
Gambar 4.8	25
Gambar 4.9	26
Gambar 4.10	27

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Sepak bola adalah olahraga yang paling populer di dunia, dengan jumlah penggemar yang diperkirakan mencapai sekitar 4 miliar orang. Angka ini menjadikannya sebagai olahraga teratas, jauh mengungguli olahraga lain seperti kriket dan hoki, yang masing-masing memiliki sekitar 2,5 miliar dan 2 miliar penggemar (Vika Azkiya Dihni, 2021). Di Indonesia, sepak bola juga sangat digemari. Menurut survei Ipsos pada September 2022, Indonesia tercatat sebagai negara dengan persentase penggemar sepak bola tertinggi di dunia, mencapai 69% dari responden yang disurvei (Annur, 2022). Survei yang dilakukan di 34 negara dengan total responden berjumlah 22.528 orang dengan rentang usia 18-74 tahun.

Kemajuan sepak bola Indonesia dalam beberapa tahun terakhir terlihat jelas melalui peningkatan kualitas kompetisi domestik, terutama di Liga 1, didukung oleh perekrutan pemain asing berpengalaman yang membantu pengembangan pemain lokal. (Asad Arifin, 2024). Program Garuda Select juga berperan penting sejak 2019, mengirim pemain U-17 ke Inggris untuk pelatihan standar Eropa. Keberhasilannya tercermin dari alumni seperti Bagus Kahfi yang kini bermain di FC Utrecht.(Arif Budi, 2022).

Salah satu tolak ukur kemajuan sepak bola Indonesia adalah masuknya Indonesia dalam kualifikasi piala dunia 2026. Namun, Meskipun pencapaian ini menunjukkan kemajuan dalam pengembangan sepak bola di Indonesia, penerapan teknologi seperti machine learning dalam sepak bola Indonesia masih belum maksimal walaupun machine learning memiliki kemampuan untuk memproses

sejumlah besar data, peningkatan performa pemain, operasi liga yang lebih efisien, menganalisis kinerja pemain serta perilaku tim lawan (*ABITA LLC & MARKETING JAPAN*, 2024).

Hal ini menjadi tantangan yang semakin nyata ketika melihat kemampuan Timnas Indonesia adalah yang terlemah dan peringkat *FIFA* yang terpaut jauh dari semua negara yang ikut serta dalam kualifikasi piala dunia 2026 di round 3 khususnya grup c. Kesenjangan kemampuan ini menjadi masalah serius karena memperkecil peluang Timnas Indonesia untuk bersaing secara kompetitif di grup dan lolos ke babak berikutnya. Selain itu, Harapan untuk lolos ke Piala Dunia, panggung sepak bola terbesar di dunia, semakin membakar semangat juang Timnas dan mengobarkan ekspektasi public (Ananto Sajatinings, 2024).

Sebagai langkah awal untuk membantu Indonesia dalam menghadapi kesenjangan kemampuan ini, penerapan teknologi seperti *machine learning* dapat menjadi solusi strategis. Teknologi ini memungkinkan analisis yang lebih mendalam, seperti mendeteksi pola permainan para pemain, evaluasi performa pemain, optimasi strategi latihan, prediksi hasil pertandingan dan sebagainya (Maulida, 2023).

Oleh karena itu, penelitian ini menggunakan pendekatan prediktif berbasis *machine learning* dengan algoritma *Support Vector Machine* karena memiliki keunggulan dalam menangani dataset yang memiliki banyak fitur, akurasi yang tinggi, serta menangani data tidak linear (Dadan Dahman W, 2021). Berdasarkan penelitian terdahulu yang dilakukan oleh (Karim et al., n.d.) yang memprediksi akurasi pertandingan liga Italia, algoritma SVM memiliki tingkat akurasi tertinggi dibanding algoritma Knn dan *Random forest*. SVM bekerja dengan cara memproses

data historis pertandingan untuk memprediksi kemungkinan hasil di masa depan. Proses ini dimulai dengan pelatihan model menggunakan data training untuk mengenali pola hubungan antara data statistik dengan posisi sementara. Setelah model memahami pola, SVM kemudian mencoba memprediksi posisi tim menggunakan data testing dengan membuat beberapa hyperplane untuk membandingkan setiap posisi dengan yang lain menggunakan metode One-Versus-Rest (OVR), sehingga dapat menentukan posisi akhir bagi masing-masing tim.

Penelitian ini bertujuan untuk memprediksi posisi akhir Timnas Indonesia di Ronde 3 Kualifikasi Piala Dunia 2026 menggunakan machine learning, sehingga memberikan gambaran sejauh mana Indonesia dapat bersaing di grupnya. Hasil prediksi ini diharapkan dapat memberikan wawasan berbasis data mengenai peluang Timnas dalam kompetisi, serta menunjukkan bagaimana inovasi teknologi dapat dimanfaatkan dalam dunia olahraga, khususnya sepak bola, untuk memberikan perkiraan yang lebih objektif terhadap hasil pertandingan.

1.2 Rumusan Masalah

Penggunaan *Machine Learning* untuk memprediksi posisi Tim Nasional Indonesia di akhir Round 3 Kualifikasi Piala Dunia Grup C masih jarang dilakukan, padahal teknik ini berpotensi memberikan prediksi yang lebih akurat dan objektif berdasarkan data statistik yang tersedia.

1.3 Batasan Masalah

1. Penelitian ini hanya akan menggunakan metode *machine learning* tertentu, seperti klasifikasi dan algoritma *Support Vector Machine*.
2. Penelitian ini hanya focus pada Kualifikasi Piala Dunia Grup C, khususnya round 3.

3. Hanya statistik yang terukur dan dapat diakses yang akan digunakan di penelitian ini, seperti penguasaan bola, jumlah tembakan, tendangan sudut dan data terkait lainnya.
4. Penelitian ini menggunakan data pertandingan yang telah selesai hingga tanggal tertentu. Kompetisi yang masih berlangsung tidak dimasukkan dalam penelitian ini karena keterbatasan waktu dan cakupan penelitian.

1.4 Tujuan Penelitian

Menerapkan metode *machine learning* untuk memprediksi posisi tim nasional Indonesia di round 3 kualifikasi piala dunia grup c berdasarkan data statistik pertandingan.

1.5 Manfaat Penelitian

1. Penelitian ini memberikan gambaran prediktif mengenai peluang Tim Nasional Indonesia dalam Kualifikasi Piala Dunia Grup C. Hasil prediksi ini dapat digunakan untuk memahami sejauh mana Indonesia dapat bersaing di grupnya serta menunjukkan bagaimana pendekatan berbasis data dapat diterapkan dalam olahraga, khususnya sepak bola.
2. Dengan memperkenalkan konsep-konsep analisis statistik dan *Machine Learning* kepada masyarakat, penelitian ini bisa membuka wawasan baru mengenai bagaimana teknologi dapat diterapkan dalam olahraga dan bagaimana statistik berperan dalam memprediksi hasil kompetisi.
3. Penelitian ini dapat mendorong lebih banyak pelatih, manajer tim, dan organisasi olahraga untuk mengadopsi teknologi dan analisis data dalam pengambilan keputusan yang lebih baik dalam strategi olahraga, termasuk untuk tim nasional Indonesia.

BAB II

LANDASAN TEORI

2.1 Sepak Bola

Menurut Komaruddin sepak bola adalah suatu kegiatan fisik yang didalamnya memiliki banyak sekali pergerakan serta bias dilihat dari taksonomi gerak umum (Harris, 2021). Tujuan utama dari permainan ini adalah mencetak gol dengan memasukkan bola ke gawang lawan. Sepak bola melibatkan berbagai aspek teknik, taktik, dan fisik yang harus dikuasai oleh pemain untuk mencapai performa terbaik. Dalam konteks ini, pemahaman yang mendalam tentang permainan tidak hanya bergantung pada keterampilan individu, tetapi juga pada analisis data yang dapat membantu tim meraih kesuksesan.

Seiring dengan perkembangan teknologi dan analisis data, statistik dalam sepak bola telah menjadi alat penting untuk menganalisis performa tim dan pemain. Penggunaan statistik memungkinkan pelatih dan manajer untuk membuat keputusan yang lebih baik berdasarkan data objektif. Dengan semakin banyaknya informasi yang tersedia, statistik kini menjadi bagian integral dari strategi tim, membantu mereka memahami dinamika permainan dan mengidentifikasi area yang perlu diperbaiki. Hal ini menunjukkan bahwa penguasaan data dapat memberikan keunggulan kompetitif dalam dunia sepak bola modern (Maulida, 2023).

Berbagai jenis data statistik digunakan untuk menganalisis performa tim dan pemain. Salah satu statistik utama adalah jumlah gol yang dicetak, yang mencerminkan efektivitas serangan suatu tim. Selain itu, assist yang menunjukkan kemampuan seorang pemain dalam menciptakan peluang bagi rekan satu tim juga menjadi statistik penting dalam menilai kontribusi pemain. Di samping itu,

penguasaan bola (possession) memberikan gambaran tentang kontrol permainan, sementara tembakan (shots) menunjukkan seberapa agresif suatu tim dalam menyerang gawang lawan (SKRIPSI FULL TEKS, n.d.).

Statistik lainnya, seperti pelanggaran (*fouls*), kartu kuning dan merah, serta statistik pertahanan seperti tekel sukses dan intersep, memberikan wawasan tambahan tentang disiplin dan efektivitas pertahanan tim. Kiper juga memiliki statistik khusus, termasuk jumlah penyelamatan (*saves*) dan gol yang kebobolan, yang sangat penting untuk mengevaluasi performa mereka di lapangan. Semua data ini saling terkait dan memberikan gambaran menyeluruh tentang bagaimana suatu tim beroperasi di dalam pertandingan.

Dengan demikian, statistik dalam sepak bola memainkan peran penting dalam analisis performa tim dan individu. Dengan menggunakan data statistik secara efektif, pelatih dapat mengevaluasi kekuatan dan kelemahan tim mereka serta merumuskan strategi permainan yang lebih efektif. Seiring dengan kemajuan teknologi, penggunaan analisis data dalam sepak bola akan terus berkembang, memberikan wawasan yang lebih mendalam untuk meningkatkan performa di lapangan. Ini menunjukkan bahwa penggabungan teori sepak bola dengan pendekatan berbasis data adalah kunci untuk mencapai kesuksesan di era modern ini.

2.2 Machine Learning

Menurut Ambler (2024) *Machine learning* adalah cabang kecerdasan buatan yang berfokus pada pembangunan sistem yang mampu belajar dari data, mengidentifikasi pola, dan membuat keputusan dengan campur tangan manusia yang minimal. Dalam analisis olahraga, teknologi ini menganalisis sejumlah besar

data performa, memberikan wawasan yang sebelumnya tidak dapat diperoleh melalui metode tradisional.

Dalam sepak bola, *machine learning* digunakan untuk meningkatkan performa pemain dan tim. Tim menggunakan algoritma *machine learning* untuk menganalisis data pertandingan, melacak pergerakan pemain, umpan, dan tembakan untuk mendapatkan wawasan tentang taktik dan performa. Pendekatan berbasis data ini membantu pelatih mengembangkan rencana permainan yang lebih efektif dan mengoptimalkan posisi pemain di lapangan.

Machine learning juga memainkan peran penting dalam perekrutan pemain. Dengan mengevaluasi sejumlah besar data performa dari para pemain di seluruh dunia, algoritma dapat mengidentifikasi bakat potensial yang sesuai dengan kebutuhan khusus suatu tim. Analisis objektif ini mengurangi risiko yang terkait dengan transfer pemain dan membantu tim membangun skuad yang lebih kuat.

Machine learning membantu dalam pengambilan keputusan secara real-time selama pertandingan. Pelatih dapat menerima umpan data langsung yang dianalisis oleh algoritma, yang memberikan wawasan tentang strategi lawan dan menyarankan penyesuaian taktis. Kemampuan untuk beradaptasi dengan cepat ini dapat menjadi faktor penentu dalam pertandingan berisiko tinggi, yang memberi tim keunggulan kompetitif (Ambler, 2024).

2.3 SVM (*Support Vector Machine*)

Support Vector Machine (SVM) merupakan salah satu metode dalam supervised learning yang biasanya digunakan untuk klasifikasi seperti (*Support Vector Classification*) dan regresi (*Support Vector Regression*). Dalam pemodelan klasifikasi, SVM memiliki konsep yang lebih matang dan lebih jelas secara

matematis dibandingkan dengan teknik-teknik klasifikasi lainnya. SVM juga dapat mengatasi masalah klasifikasi dan regresi dengan linier maupun non linier (Samsudiney, 2019).

2.4 Python

Python merupakan bahasa pemrograman tingkat tinggi yang yang dibuat oleh Guido Van Rossum dan dirilis pada tahun 1991 *Python* juga merupakan bahasa sangat populer belakangan ini. Selain itu *Python* juga merupakan bahasa pemrograman yang multi fungsi contohnya *Python* dapat digunakan untuk *Machine learning* dan *Deep Learning*. *Python* dipilih sebagai penelitian karena *Python* memiliki penulisan sintaksis yang mudah selain itu *Python* juga memiliki *library* yang lengkap dan memiliki dukungan komunitas yang kuat karena *Python* bersifat open source. Untuk menuliskan source code *Python* anda dapat menggunakan *IDE* seperti vs code, *sublime text*, *PyCharm* atau anda juga dapat menggunakan *IDE* online seperti Jupyter notebook dan google colab (Riziq sirfatullah Alfarizi et al., 2023).

2.5 Classification Report

Classification Report adalah alat evaluasi dalam *machine learning* untuk meringkas kinerja model klasifikasi. Beberapa metrik utama yang digunakan seperti, Precision, Recall, F1-Score, dan Support, untuk setiap kelas dalam dataset (*Classification_report*, 2025).

- *Precision*

Merupakan rasio antara kelas yang diprediksi dengan benar dengan keseluruhan kelas yang diprediksi.

$$Precision = \frac{TP}{TP+FP} \quad \begin{array}{l} TP = \text{True Positives} \\ FP = \text{False Positive} \end{array} \quad (1)$$

- *Recall*

Merupakan rasio kelas yang diprediksi dengan benar terhadap semua data di kelas sebenarnya.

$$Recall = \frac{TP}{TP+FN} \quad \begin{array}{l} TP = \text{True Positives} \\ FN = \text{False Negatives} \end{array} \quad (2)$$

- *F1-Score*

Merupakan rata-rata harmonis antara skor Precision dan Recall.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3)$$

- *Support*

Jumlah sampel sebenarnya di setiap kelas.

2.6 Penelitian Terkait

Penelitian sebelumnya yang dilakukan (Zein & Gunawan, 2022) membahas tentang hasil prediksi pertandingan *FIFA World Cup 2022* dengan bantuan *library Scikit-Learn* pada *Python* menggunakan berbagai algoritma *machine learning* seperti *Logistic Regression*, *K-Nearest Neighbors*, *Naïve Bayes*, *Support Vector Machine (SVM)*, *Neural Network*, dan *Random forest*. Penelitian ini menggunakan data historis *FIFA World Cup* dari tahun 2006 hingga 2018, 9 statistic performa tim pada permainan eFootball 2022, serta jadwal *FIFA World Cup 2022*. Hasil penelitian menunjukkan bahwa SVM memiliki akurasi rata-rata tertinggi sebesar 59.37%, sedangkan *Neural Network* memberikan hasil terbaik pada edisi tertentu (2006 dan 2018). Prediksi untuk *FIFA World Cup Qatar 2022* menunjukkan bahwa Jerman memiliki peluang besar menjadi juara, dengan algoritma *Neural Network*

digunakan untuk prediksi berdasarkan histori 12 tahun lalu dan SVM untuk histori lebih dekat atau data baru.

Pada penelitian oleh (Qomariah & Rangan, 2020) membahas tentang klasifikasi tipe gelandang sepak bola berdasarkan kemampuan pemain menggunakan algoritma *Naïve Bayes*. Data yang digunakan terdiri dari 101 pemain aktif dengan posisi CMF (*central midfielder*), AMF (*attacking midfielder*), dan DMF (*defensive midfielder*). Penelitian ini menggunakan alat Weka untuk membangun model dengan empat skenario pembagian data latih dan uji, termasuk metode cross-validation 10-*folds*. Hasil penelitian menunjukkan bahwa akurasi terbaik sebesar 82.5% diperoleh pada skenario pembagian data 60/40.

Pada penelitian oleh (Karim et al., n.d.) membahas tentang perbandingan metode *Random forest*, *K-Nearest Neighbor* (KNN), dan *Support Vector Machine* (SVM) dalam memprediksi akurasi pertandingan Liga Italia musim 2020/2021. Tujuan utama penelitian ini adalah untuk mengevaluasi efektivitas masing-masing algoritma dalam memprediksi hasil pertandingan berdasarkan data yang tersedia. Dalam pengujian menggunakan skema pembagian data 80% untuk pelatihan dan 20% untuk pengujian, hasil menunjukkan bahwa algoritma SVM memiliki tingkat akurasi tertinggi sebesar 64%, diikuti oleh *Random forest* dengan akurasi 62%, dan KNN yang memperoleh akurasi terendah sebesar 57%. Hasil ini menunjukkan bahwa SVM lebih efektif dalam memprediksi hasil pertandingan dibandingkan dengan kedua algoritma lainnya.

Pada penelitian oleh (Pratama & Nurhasanah, 2020) membahas tentang penentuan starting eleven pada sepak bola menggunakan metode *Profile Matching* dan *Naïve bayes*, bertujuan untuk meningkatkan objektivitas pemilihan pemain oleh

pelatih. Sistem yang dikembangkan menerima input berupa nilai atribut kemampuan dan kondisi pemain, di mana proses *Profile Matching* digunakan untuk menentukan posisi ideal, sementara *Naïve bayes* digunakan untuk memilih pemain yang cocok sebagai starting eleven. Hasil pengujian menunjukkan akurasi sistem mencapai 65%, namun peningkatan rata-rata rating tim setelah penerapan sistem hanya sebesar 0,98%, mengindikasikan bahwa meskipun sistem ini membantu dalam pemilihan pemain, faktor lain juga berperan dalam performa tim secara keseluruhan.

Pada penelitian oleh (Liazizah Putri & Riyardi, n.d.) membahas prediksi hasil pertandingan Piala Dunia 2022 menggunakan metode *Decision Tree* dan algoritma *CART*, dengan tujuan untuk menunjukkan kinerja algoritma dalam memprediksi hasil pertandingan serta pentingnya prediksi dalam menyusun strategi tim dalam olahraga sepakbola. Hasil penelitian menunjukkan bahwa algoritma *CART* memiliki akurasi sebesar 0.6659, yang menandakan kinerjanya yang baik dalam memprediksi hasil pertandingan tanpa tanda *overfitting*.

Tabel 2.1 Penelitian Terkait

No	Nama Peneliti (Tahun)	Penelitian	Hasil
1	Syahrul Zein, Gani Gunawan (2022)	Prediksi Hasil <i>FIFA</i> World Cup Qatar 2022 Menggunakan <i>Machine learning</i> dengan <i>Python</i>	Hasil penelitian menunjukkan bahwa SVM memiliki akurasi rata-rata tertinggi sebesar 59.37%, sedangkan <i>Neural Network</i> memberikan hasil terbaik pada edisi tertentu (2006 dan 2018). Prediksi untuk <i>FIFA</i> World Cup Qatar 2022 menunjukkan bahwa Jerman memiliki peluang besar menjadi juara, dengan algoritma <i>Neural Network</i> digunakan untuk prediksi berdasarkan histori 12 tahun lalu dan SVM untuk histori lebih dekat atau data baru.
2	Ni Wayan Wisswani, Tien Rahayu Tulili, Muhammad	Klasifikasi Tipe Gelandang Sepak Bola Berdasarkan Data Kemampuan	Hasil penelitian menunjukkan bahwa akurasi terbaik sebesar 82.5%

	Farman Andrijasa, M. Rangga Satria N, Siti Jumani, Eko Wahyudi (2020).	Menggunakan Metode <i>Naïve Bayes</i>	diperoleh pada skenario pembagian data 60/40
3	Ahmad Assril Karim, Muhammad Ary Prasetyo, Muhammad Rohid Saputro (2023)	Perbandingan Metode <i>Random forest</i> , <i>K-Nearest Neighbor</i> , dan SVM Dalam Prediksi Akurasi Pertandingan Liga Italia	Dalam pengujian menggunakan skema pembagian data 80% untuk pelatihan dan 20% untuk pengujian, hasil menunjukkan bahwa algoritma SVM sebesar 64%, diikuti oleh <i>Random forest</i> dengan akurasi 62%, dan KNN yang memperoleh akurasi terendah sebesar 57%. Hasil ini menunjukkan bahwa SVM lebih efektif dalam memprediksi hasil pertandingan dibandingkan dengan kedua algoritma lainnya
4	Fabio Fahri Pratama, Youllia	Penggunaan Metode <i>Profile Matching</i> Dan	Hasil pengujian menunjukkan akurasi sistem

	Indrawaty Nurhasanah (2020)	<i>Naïve bayes</i> Untuk Menentukan Starting <i>Eleven</i> Pada Sepak Bola	mencapai 65%, namun peningkatan rata-rata rating tim setelah penerapan sistem hanya sebesar 0,98%, mengindikasikan bahwa meskipun sistem ini membantu dalam pemilihan pemain, faktor lain juga berperan dalam performa tim secara keseluruhan.
5	Prihananti Liazizah Putri, Agung Riyardi (2022)	Analisis Prediksi Piala Dunia 2022 Dengan Menggunakan Metode <i>Decision Tree</i> Dan Algoritma Cart	Hasil penelitian menunjukkan bahwa algoritma CART memiliki akurasi sebesar 0.6659, yang menandakan kinerjanya yang baik dalam memprediksi hasil pertandingan tanpa tanda <i>overfitting</i> .

BAB III

METODOLOGI PENELITIAN

3.1 Pendekatan Penelitian

Dalam penelitian ini, pendekatan yang digunakan adalah pendekatan prediktif dengan menggunakan model klasifikasi algoritma SVM untuk memprediksi hasil pertandingan sepak bola. Algoritma ini dipilih karena memiliki kemampuan yang unggul dalam menangani data statistik seperti penguasaan bola, tendangan sudut dan pelanggaran untuk menghasilkan prediksi peringkat tim. Metode ini juga memiliki keunggulan dalam menangani dataset yang memiliki banyak fitur, akurasi yang tinggi, serta ketahanannya terhadap *overfitting*.

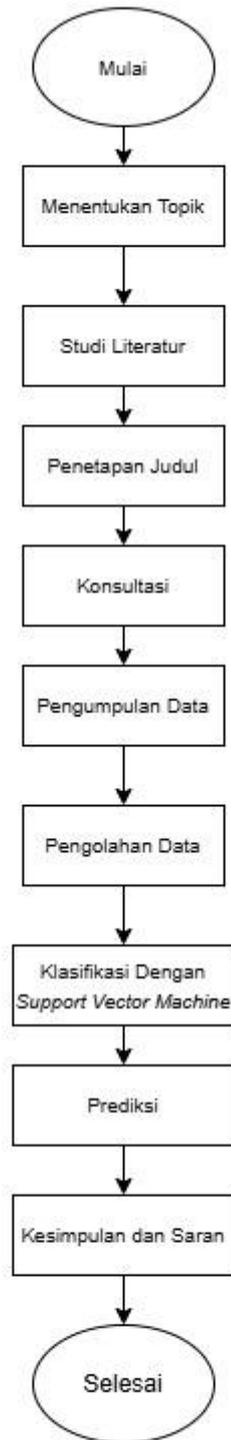
3.2 Teknik Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data sekunder dengan memanfaatkan sumber dari livescore, yang menyediakan informasi detail tentang pertandingan sepak bola seluruh dunia. Metode pengumpulan data yang digunakan adalah dokumentasi secara manual dan memasukkannya ke Microsoft Excel. Data yang dikumpulkan meliputi statistik pertandingan, seperti :

1. Match
2. Negara
3. Lawan
4. Penguasaan bola
5. Tembakan ke gawang
6. Tembakan tepat sasaran
7. Tembakan melenceng dari sasaran
8. Tembakan yang diblokir

9. Penyelamatan penjaga
10. Tendangan bebas
11. Offside
12. Pelanggaran
13. Kartu kuning
14. Kartu merah
15. Lemparan kedalam
16. Umpan
17. Intersepsi
18. Posisi sementara

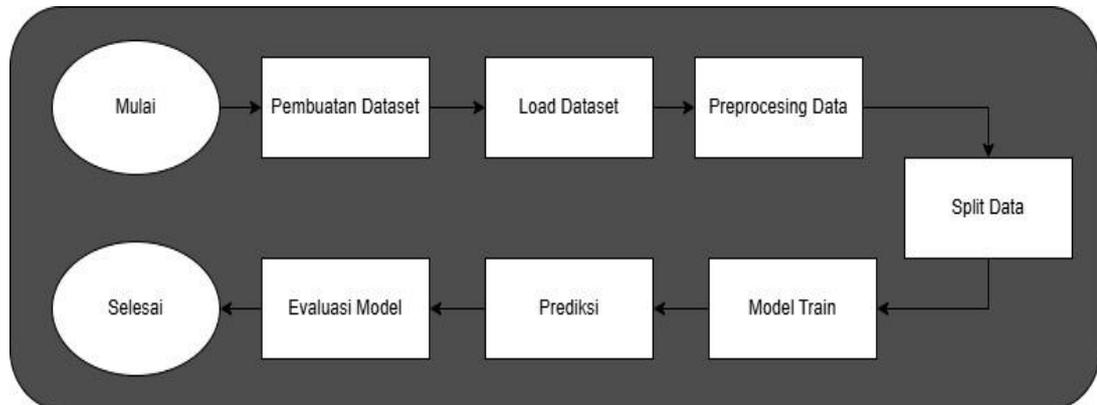
3.3 Tahapan Penelitian



Gambar 3.1 Tahapan Penelitian

1. Tahapan pertama menentukan topik penelitian yang akan diambil. Dimana penelitian ini mengambil topik *Machine learning* dan mencari permasalahan yang bertujuan menentukan arah penelitian yang akan dilakukan.
2. Tahapan kedua adalah studi literatur. Dimana peneliti mencari referensi tentang masalah penelitian dari berbagai sumber, seperti penelitian terdahulu, jurnal, artikel, berita terkait dengan masalah penelitian.
3. Tahapan ketiga yaitu penetapan judul. Dimana peneliti mengajukan judul penelitian dan disetujui oleh ketua program studi.
4. Tahapan keempat adalah konsultasi. Dimana setelah judul telah ditetapkan kemudian peneliti berdiskusi dengan dosen pembimbing yang telah ditetapkan terkait penelitian.
5. Tahapan kelima adalah pengumpulan data. Mengumpulkan data yang relevan untuk penelitian dari sumber terpercaya.
6. Tahapan keenam yaitu pengolahan data. Data yang telah dikumpulkan akan diolah menggunakan algoritma SVM dengan model klasifikasi.
7. Tahapan ketujuh adalah prediksi. Menampilkan hasil prediksi dari data yang telah diolah menggunakan algoritma SVM dengan model klasifikasi
8. Tahapan kedelapan yaitu kesimpulan dan saran. Kesimpulan diambil dari keseluruhan penelitian dan memberikan saran untuk penelitian selanjutnya.

3.4 Alur Pengolahan Data



Gambar 3.2 Alur Pengolahan Data

1. Hal yang dilakukan pertama kali yaitu pembuatan dataset. Dataset yang digunakan pada penelitian ini diambil dari website Livescore.in yang dipindahkan ke excel secara manual.
2. Selanjutnya setelah dataset telah dibuat dan tersimpan di file excel, dataset akan dibaca. File ini berisi data statistik pertandingan yang mencakup berbagai informasi seperti penguasaan bola, tendangan sudut, tendangan bebas dan data terkait statistik lainnya.
3. Langkah selanjutnya adalah preprocessing data yaitu data yang tidak relevan atau tidak diperlukan akan di hapus. Kemudian data dibagi menjadi 2 bagian yaitu fitur (x) dan target (y). Fitur adalah variabel yang digunakan untuk melakukan prediksi (statistik pertandingan), sedangkan target adalah posisi sementara yang ingin diprediksi
4. Setelah data di preprocessing, data akan dibagi dengan presentasi 80% untuk data training dan 20% untuk data testing. Data training digunakan untuk melatih model agar dapat mengenali pola dan hubungan dalam data, sedangkan data testing digunakan untuk mengevaluasi kinerja model setelah dilatih.

5. Setelah pembagian data, kemudian model dilatih menggunakan data training. Pada penelitian ini model yang digunakan adalah SVM, merupakan algoritma yang sering digunakan untuk klasifikasi dan regresi.
6. Setelah model dilatih langkah selanjutnya adalah melakukan prediksi pada data testing menggunakan model yang sudah dilatih. Prediksi ini dipilih berdasarkan probabilitas tertinggi untuk menentukan posisi sementara yang diperkirakan.
7. Setelah prediksi dilakukan selanjutnya adalah tahap evaluasi. Tahap evaluasi dilakukan untuk menilai seberapa baik model dalam memprediksi data uji. Metrik utama yang digunakan adalah *Precision*, *Recall*, *F1-Score*, dan *Support*. *Precision* mengukur rasio antara prediksi benar untuk suatu kelas dengan total prediksi untuk kelas tersebut, *Recall* mengukur rasio prediksi benar untuk suatu kelas dengan total data aktual pada kelas tersebut. *F1-Score* merupakan rata-rata harmonis antara *Precision* dan *Recall*, memberikan keseimbangan antara kedua metrik tersebut, *Support* yaitu jumlah sampel aktual yang terdapat pada setiap kelas.

BAB IV HASIL DAN PEMBAHASAN

4.1. Hasil Pengumpulan Data

Tahap awal penelitian dimulai dengan pengumpulan data melalui website *Livescore.in*. Data yang dikumpulkan mencakup statistik pertandingan seperti penguasaan bola, tendangan sudut, tendangan bebas dan informasi terkait statistik lainnya dalam beberapa pertandingan terakhir. Proses pengumpulan dilakukan secara manual dengan memindahkan data ke dalam file excel untuk memudahkan pengolahan data.

match	negara	lawan	penguasaan bola	tembakan ke gawang	tembakan tepat sasaran	tembakan melenceng dari sasaran	sudut	penyelamatan penjaga	tendangan bebas	offside	pelanggaran	kartu kuning	kartu merah	menyerah	lulus	intersepsi	posisi sementara	
1	indonesia	china	76	14	6	5	3	6	1	13	1	8	2	0	35	83	11	4
2	indonesia	jepang	33	9	3	3	3	4	4	8	2	9	0	1	25	78	14	4
3	indonesia	arab	23	13	6	4	3	4	3	4	5	16	5	1	14	58	9	4
4	indonesia	australia	60	11	4	3	4	7	2	11	1	13	0	0	26	86	5	4
5	indonesia	bahrain	42	7	3	4	0	4	1	8	4	14	2	0	26	80	9	4
1	arab	bahrain	67	14	4	6	4	8	2	14	5	9	1	0	31	79	9	3
2	arab	australia	58	4	1	1	2	3	2	8	5	9	2	0	0	77	11	3
3	arab	indonesia	77	25	3	12	10	12	4	16	1	4	3	0	25	88	12	3
4	arab	china	80	19	7	8	4	5	0	9	5	4	0	0	24	91	5	3
5	arab	jepang	22	1	0	1	0	0	2	10	0	3	0	0	18	52	10	3
1	australia	jepang	37	2	1	1	0	0	3	12	0	10	4	0	35	76	24	2
2	australia	arab	42	14	2	5	7	3	1	8	7	11	2	0	0	77	11	2
3	australia	bahrain	59	8	4	3	1	6	1	13	5	11	1	0	27	81	10	2
4	australia	indonesia	40	9	7	1	1	5	3	13	0	11	0	0	20	82	7	2

4.1 Hasil Pengumpulan Data

4.2. Import Library

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.svm import SVC
from sklearn.metrics import classification_report
import numpy as np
```

4.2 Gambar Kode Import Library

Dari gambar 4.2 diatas terlihat ada beberapa *library* yang digunakan yaitu pandas digunakan untuk membaca dan mengelola data dalam bentuk tabel agar lebih mudah diproses. Kemudian, `train_test_split` dari `sklearn.model_selection` digunakan untuk membagi dataset menjadi data latih dan data uji, sehingga model bisa belajar dari sebagian data dan diuji dengan data yang belum pernah dilihat sebelumnya. `SVC` dari `sklearn.svm` adalah algoritma *Support Vector Classifier* yang digunakan untuk melakukan klasifikasi dengan mencari hyperplane terbaik yang memisahkan kelas-kelas dalam data. Setelah model dilatih, `classification_report` dari `sklearn.metrics` digunakan untuk mengevaluasi performanya dengan memberikan metrik seperti precision, recall, dan f1-score. Kemudian `numpy` digunakan untuk manipulasi array dan perhitungan numerik.

4.3. Import Dataset

```
file_path = '4hingga8.xlsx'
data = pd.read_excel(file_path)
```

4.3 Gambar Kode Import Dataset

Kode dalam gambar 4.3 menunjukkan bahwa file Excel yang lokasinya ditentukan oleh `file_path` akan dibaca, kemudian isinya disimpan dalam variabel `data` untuk digunakan dalam program.

4.4. Pre-Processing Data

```
# Preprocessing data
data_cleaned = data.drop(columns=['match', 'lawan'])

# Pisahkan fitur dan target
X = data_cleaned.drop(columns=['posisi sementara'])
y = data['posisi sementara']
```

4.4 Gambar Kode Pre-Processing Data

Kode yang ada di gambar 4.4 menunjukkan proses penghapusan kolom yang tidak berpengaruh terhadap prediksi, yaitu match dan lawan, sehingga hanya kolom yang diperlukan yang tersisa dalam dataset. Setelah itu, kode ini memisahkan fitur (X) dan target (Y), di mana x berisi semua fitur dalam dataset kecuali kolom yang sudah dihapus sebelumnya serta posisi sementara, karena kolom tersebut adalah target yang ingin diprediksi. Sementara itu, y berisi nilai target yang akan diprediksi, yaitu posisi sementara.

4.5. Split Data

```
# Split data untuk training dan testing
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=52, stratify=y)
print(f'Panjang X_train: {len(X_train)}')
print(f'Panjang X_test: {len(X_test)}')
```

4.5 Gambar Kode Split Data

Kode ini menjelaskan pembagian data menjadi data latih dan data uji dengan rasio 80% untuk data latih dan 20% untuk data uji. Kemudian, random_state digunakan untuk memastikan pembagian data tetap sama setiap kali kode dijalankan, sementara stratify menjaga proporsi kelas dalam data latih dan uji tetap seimbang. Setelah itu, kode print(f'Panjang X_train;{len(X_train)}') dan print(f'Panjang X_test;{len(X_test)}') digunakan untuk menampilkan jumlah data latih dan data uji yang telah dibagi sebelumnya.

4.6. Modeling Svm

```
# Model SVM dengan probabilitas
svm_model = SVC(kernel='linear', C=0.1, probability=True, random_state=42)
svm_model.fit(X_train, y_train)
```

4.6 Gambar Kode Modeling Svm

Kode yang ada dalam gambar 4.6 menjelaskan proses pembangunan model *Support Vector Machine (SVM)* untuk klasifikasi. Model dibuat menggunakan *SVC* dengan kernel linear (`kernel='linear'`) dan parameter `C=0.1`, yang berfungsi untuk mengontrol keseimbangan antara akurasi klasifikasi dan margin pemisahan data. Selain itu, parameter `probability=True` diaktifkan agar model dapat menghitung probabilitas dari setiap kelas dalam prediksinya. Setelah model dibuat, `svm` digunakan untuk melatih model dengan data latih, sehingga model dapat mengenali pola dalam data dan siap digunakan untuk melakukan prediksi pada data uji

4.7. Hasil Prediksi

```
y_pred = svm_model.predict(X_test)
# Prediksi probabilitas pada data uji
y_pred_proba = svm_model.predict_proba(X_test)

# Ambil 2 prediksi dengan probabilitas tertinggi
top_pred_idx = np.argsort(y_pred_proba, axis=1)[:, -2:]
top_pred_proba = np.take_along_axis(y_pred_proba, top_pred_idx, axis=1)

# Urutkan agar prediksi tertinggi ada di depan
sorted_indices = np.argsort(-top_pred_proba, axis=1)
top_pred_idx = np.take_along_axis(top_pred_idx, sorted_indices, axis=1)
top_pred_proba = np.take_along_axis(top_pred_proba, sorted_indices, axis=1)

# Kembalikan ke posisi asli (1-based)
top_pred_idx += 1

# Buat DataFrame hasil
result_df = pd.DataFrame({
    'Sementara': y_test.values + 1,
    'Prediksi 1': top_pred_idx[:, 0],
    'Prediksi 2': top_pred_idx[:, 1],
    'Prob 1 (%)': (top_pred_proba[:, 0] * 100).round(2),
    'Prob 2 (%)': (top_pred_proba[:, 1] * 100).round(2),
})
```

4.7 Gambar Kode Hasil Prediksi

Kode di atas digunakan untuk memprediksi label dari data uji berdasarkan pola yang telah dipelajari, dan hasil prediksi tersebut disimpan dalam variabel `y_pred` kemudian prediksi probabilitas menggunakan model *SVM* dengan `predict_proba(X_test)`, lalu mengambil dua prediksi dengan probabilitas tertinggi menggunakan `np.argsort(y_pred_proba, axis=1)[-2:]`, yang mengurutkan indeks dari probabilitas terkecil ke terbesar dan memilih dua yang terakhir. Setelah itu, `np.take_along_axis` digunakan untuk mendapatkan nilai probabilitas dari indeks tersebut. Agar prediksi dengan probabilitas tertinggi berada di depan, kode `np.argsort(-top_pred_proba, axis=1)` mengurutkan dalam urutan descending, lalu `np.take_along_axis` memastikan indeks dan probabilitas disesuaikan dengan urutan baru. Setelah proses sorting selesai, `top_pred_idx += 1` mengubah indeks prediksi dari 0-based ke 1-based agar sesuai dengan skala asli. Terakhir, hasil prediksi disusun dalam DataFrame dengan kolom Sementara (nilai sebenarnya), Prediksi 1 dan Prediksi 2 (dua prediksi dengan probabilitas tertinggi), serta Prob 1 (%) dan Prob 2 (%) yang menunjukkan persentase probabilitas masing-masing prediksi.

Hasil Prediksi dan Probabilitas:

	Sementara	Prediksi 1	Prediksi 2	Prob 1 (%)	Prob 2 (%)
0	1	1	1	35.67	20.07
1	2	3	4	20.42	19.27
2	3	2	3	34.31	22.02
3	4	2	4	19.73	17.43
4	5	5	3	23.09	19.89
5	6	5	4	24.95	24.21

4.8 Gambar Hasil Prediksi

Pada gambar 4.8 hasil prediksi menunjukkan :

- a) Posisi 1 diperkirakan tetap di posisi 1 dengan probabilitas 35.67% dan 20.07%.
- b) Posisi 2 diperkirakan akan turun ke posisi 3 dengan probabilitas 20.42% atau turun ke posisi 4 dengan probabilitas 19.27%.

- c) Posisi 3 diperkirakan akan naik ke posisi 2 dengan probabilitas 34.31% atau menetap di posisi 3 dengan probabilitas 22.02%.
- d) Posisi 4 diperkirakan akan naik ke posisi 2 dengan probabilitas 19.73% atau menetap di posisi 4 dengan probabilitas 17.43%.
- e) Posisi 5 diperkirakan akan menetap di posisi 5 dengan probabilitas 23.09% atau turun ke posisi 6 dengan probabilitas 19.89%.
- f) Posisi 6 diperkirakan akan naik ke posisi 5 dengan probabilitas 24.95% atau naik ke posisi 4 dengan probabilitas 24.21%.

4.8. Evaluasi Model

```
class_report = classification_report(y_test, y_pred, target_names=[f'Posisi {i+1}' for i in range(len(np.unique(y))])
# Tampilkan laporan evaluasi dan hasil prediksi
print("Laporan Evaluasi (Precision, Recall, F1-score):")
print(class_report)
```

4.9 Gambar Kode Evaluasi Model

Kode dari gambar 4.9 di atas menggunakan fungsi `classification_report` dari library `scikit-learn` untuk menghasilkan laporan evaluasi model klasifikasi. Laporan ini berisi metrik-metrik evaluasi penting seperti precision, recall, dan F1-score untuk setiap kelas dalam dataset yang diuji, di mana nama-nama kelas diberi nama "Posisi 1", "Posisi 2", dan seterusnya sesuai dengan jumlah kelas unik dalam data target (y).

Berikut hasil dari evaluasi untuk hasil prediksi

Laporan Evaluasi (Precision, Recall, F1-score):				
	precision	recall	f1-score	support
Posisi 1	0.50	1.00	0.67	1
Posisi 2	0.50	1.00	0.67	1
Posisi 3	1.00	1.00	1.00	1
Posisi 4	0.00	0.00	0.00	1
Posisi 5	1.00	1.00	1.00	1
Posisi 6	0.00	0.00	0.00	1
accuracy			0.67	6
macro avg	0.50	0.67	0.56	6
weighted avg	0.50	0.67	0.56	6

4.10 Gambar Laporan Evaluasi Model

Pada gambar 4.10 laporan evaluasi menunjukkan :

1. Precision : Mengukur seberapa akurat model dalam memberikan prediksi untuk setiap posisi.
 - a) Posisi 1 memiliki precision 0.5, artinya setiap kali model memprediksi posisi 1, hasilnya hanya 50% yang benar.
 - b) Posisi 2 memiliki precision 0.5, artinya setiap kali model memprediksi posisi 2, hasilnya hanya 50% yang benar.
 - c) Posisi 3 memiliki precision 1.00, artinya setiap kali model memprediksi posisi 3, hasilnya selalu benar.
 - d) Posisi 4 memiliki precision 0.00, artinya setiap kali model memprediksi posisi 4, hasilnya selalu salah.
 - e) Posisi 5 memiliki precision 1.00, artinya setiap kali model memprediksi posisi 5, hasilnya selalu benar.
 - f) Posisi 6 memiliki precision 0.00, artinya setiap kali model memprediksi posisi 6, hasilnya selalu salah.

2. Recall : Mengukur seberapa banyak data actual yang benar-benar terdeteksi oleh model.
 - a) Posisi 1 memiliki recall 1.00, berarti model selalu berhasil mendeteksi posisi 1 dengan benar.
 - b) Posisi 2 memiliki recall 1.00, berarti model selalu berhasil mendeteksi posisi 2 dengan benar.
 - c) Posisi 3 memiliki recall 1.00, berarti model selalu berhasil mendeteksi posisi 3 dengan benar.
 - d) Posisi 4 memiliki recall 0.00, berarti model tidak pernah mendeteksi posisi 4 dengan benar.
 - e) Posisi 5 memiliki recall 1.00, berarti model selalu berhasil mengenali posisi 5.
 - f) Posisi 6 memiliki recall 0.00, berarti model tidak pernah mendeteksi posisi 6 dengan benar.
3. F1-Score: Rata-rata harmonis dari precision dan recall. Jika precision dan recall tinggi, maka F1-score juga tinggi.
 - a) Posisi 1 memiliki F1-score 0.67, menunjukkan bahwa performa model sedang dalam memprediksi posisi 1.
 - b) Posisi 2 memiliki F1-score 0.67, menunjukkan bahwa performa model sedang dalam memprediksi posisi 2.
 - c) Posisi 3 memiliki F1-score 1.00, menunjukkan bahwa model sangat akurat dalam memprediksi posisi 3.
 - d) Posisi 4 memiliki F1-score 0.00, menunjukkan bahwa model sangat buruk dalam memprediksi posisi 4.

- e) Posisi 5 memiliki F1-score 1.00, menunjukkan bahwa model sangat akurat dalam memprediksi posisi 5.
 - f) Posisi 6 memiliki F1-score 0.00, menunjukkan bahwa model sangat buruk dalam memprediksi posisi 6.
4. Support: Jumlah sampel untuk setiap posisi dalam data uji. Semua posisi memiliki support 1, artinya hanya ada satu contoh untuk setiap posisi.
 5. Accuracy : Persentase total prediksi yang benar, yaitu 67% (0.67) dari 6 data uji.
 6. Macro Avg : Rata-rata precision, recall, dan F1-score tanpa mempertimbangkan jumlah sampel.
 7. Weighted Avg : Rata-rata precision, recall, dan F1-score yang mempertimbangkan jumlah sampel untuk setiap kelas.

BAB V

PENUTUP

5.1 Kesimpulan

Dari penelitian yang telah dilakukan dapat disimpulkan :

1. Hasil prediksi menunjukkan bahwa posisi Timnas Indonesia akan naik ke posisi 2 dengan probabilitas 19.73% atau menetap di posisi 4 dengan probabilitas 17.43%.
2. Penelitian ini berhasil mengimplementasikan metode *Machine Learning* dengan Support Vector Machine (SVM) untuk memprediksi posisi akhir Tim Nasional Indonesia di Round 3 Kualifikasi Piala Dunia Grup C.
3. Model prediksi yang dikembangkan memiliki tingkat akurasi sebesar 67%, menunjukkan bahwa model sudah mampu mengklasifikasikan sebagian data dengan benar, namun masih memerlukan pengembangan untuk meningkatkan akurasi.

5.2 Saran

Adapun saran yang dapat diberikan oleh penulis untuk pengembang dan penelitian selanjutnya adalah sebagai berikut:

1. Meskipun penelitian ini hanya berfokus pada prediksi posisi akhir tim, pendekatan *machine learning* juga dapat diterapkan dalam penelitian lain untuk menganalisis strategi pertandingan, seperti pola permainan lawan atau efektivitas formasi tim.
2. Sebagai langkah lanjutan, model yang dikembangkan dapat diintegrasikan ke dalam aplikasi berbasis web atau mobile untuk memudahkan aksesibilitas bagi pelatih, pemain, dan penggemar sepak bola

3. Data yang digunakan dalam penelitian ini masih terbatas pada statistik pertandingan. Penelitian selanjutnya disarankan untuk mencoba algoritma machine learning lain yang lebih sesuai untuk data terbatas untuk meningkatkan akurasi prediksi

Dengan adanya saran-saran ini, diharapkan penelitian di bidang sepak bola berbasis machine learning dapat terus berkembang dan memberikan kontribusi yang lebih besar dalam dunia olahraga, khususnya dalam mendukung tim pelatih dan manajemen dalam menyusun strategi yang lebih efektif berdasarkan analisis data statistik pertandingan.

DAFTAR PUSTAKA

- Karim, A. A., Ary Prasetyo, M., & Saputro, M. R. (n.d.). *Perbandingan Metode Random Forest, K-Nearest Neighbor, dan SVM Dalam Prediksi Akurasi Pertandingan Liga Italia* (Vol. 2). <http://www.football-data.co.uk>.
- Liazizah Putri, P., & Riyardi, A. (n.d.). *ANALISIS PREDIKSI PIALA DUNIA 2022 DENGAN MENGGUNAKAN METODE DECISION TREE DAN ALGORITMA CART*.
- Pratama, F. F., & Nurhasanah, Y. I. (2020). PENGGUNAAN METODE PROFILE MATCHING DAN NAÏVE BAYES UNTUK MENENTUKAN STARTING ELEVEN PADA SEPAK BOLA. *Jurnal Tekno Insentif*, 14(2), 59–68. <https://doi.org/10.36787/jti.v14i2.268>
- Qomariah, S., & Rangan, A. Y. (2020). PERBANDINGAN METODE SAW DAN MFEP SISTEM PENDUKUNG KEPUTUSAN PENYELEKSIAN ATLET PENCAK SILAT. *Just TI (Jurnal Sains Terapan Teknologi Informasi)*, 12(1), 8. <https://doi.org/10.46964/justti.v12i1.175>
- Riziq sirfatullah Alfarizi, M., Zidan Al-farish, M., Taufiqurrahman, M., Ardiansah, G., & Elgar, M. (2023). PENGGUNAAN PYTHON SEBAGAI BAHASA PEMROGRAMAN UNTUK MACHINE LEARNING DAN DEEP LEARNING. In *Karimah Tauhid* (Vol. 2, Issue 1). *SKRIPSI FULL TEKS*. (n.d.).
- Zein, S., & Gunawan, G. (2022). Prediksi Hasil FIFA World Cup Qatar 2022 Menggunakan Machine Learning dengan Python. *Jurnal Riset Matematika*, 153–162. <https://doi.org/10.29313/jrm.v2i2.1382>
- Vika Azkiya Dihni. (2021, October 5). *Sepak Bola Jadi Olahraga Paling Populer di Dunia*. Katadata.co.id; Databoks. <https://databoks.katadata.co.id/olahraga/statistik/1389ba0bcf80468/sepak-bola-jadi-olahraga-paling-populer-di-dunia>
- Annur, C. M. (2022, December 8). Survei Ipsos: Indonesia Punya Penggemar Sepak Bola Terbesar di Dunia. Katadata.co.id; Databoks. <https://databoks.katadata.co.id/olahraga/statistik/522de0f585f0915/survei-ipsos-indonesia-punya-penggemar-sepak-bola-terbesar-di-dunia>
- Arif Budi. (2022, January 6). Soroti Program Garuda Select, Media Inggris Singgung Target Piala Dunia Timnas Indonesia. Suara.com; Suara.com. <https://www.suara.com/bola/2022/01/06/201851/soroti-program-garuda-select-media-inggris-singgung-target-piala-dunia-timnas-indonesia>
- Evolving Indonesian Football with AI: Future Tactics and Player Discovery | ABITA LLC & MARKETING JAPAN*. (2024). ABITA LLC & MARKETING JAPAN. <https://1xmarketing.com/news/en/world-marketing-diary-240825133833/?form=MG0AV3>
- Ananto Sajatining. (2024, October 23). Gelora di Balik Garuda, Ekspektasi Masyarakat pada Timnas Indonesia di Kancah Kualifikasi Piala Dunia Halaman 1 - Kompasiana.com. KOMPASIANA; Kompasiana.com. https://www.kompasiana.com/anantosajatining3371/671938b734777c01812017b2/gelora-di-balik-garuda-ekspektasi-masyarakat-pada-timnas-indonesia-di-kancah-kualifikasi-piala-dunia?form=MG0AV3#google_vignette
- Maulida, A. F. (2023, October 17). Keunggulan Teknik Machine Learning Untuk Analisis Sepakbola. Dibimbing.id; dibimbing.id.

<https://dibimbing.id/blog/detail/keunggulan-teknik-machine-learning-untuk-analisis-sepakbola>

Dadan Dahman W. (2021, July 12). *Support Vector Machine (SVM) - Howdy Sysinfo - Medium*. Medium; Howdy Sysinfo.

<https://medium.com/sysinfo/support-vector-machine-svm-5d95a7d7a547>

Harris. (2021, November 3). *Pengertian Sepak Bola: Sejarah, Peraturan, Teknik Dasar dan Manfaat*. Best Seller Gramedia. https://www.gramedia.com/best-seller/pengertiansepakbola/?srsltid=AfmBOopF6_s_a4PGzTDY2RtL2krTQJe7YiXPMNZJB5eUOgZFHuii1jM

Ambler, W. (2024, July 2). *Machine Learning in Sports Analytics | Catapult*. Catapult. <https://www-catapult-com.translate.goog/blog/sports-analytics-machinelearning? x tr sl=en& x tr tl=id& x tr hl=id& x tr pto=sge#:~:text=Dalam%20sepak%20bola%20%2C%20pembelajaran%20mesin%20di gunakan%20untuk%20meningkatkan%20performa%20pemain,yang%20me mberi%20tim%20keunggulan%20kompetitif.>

Samsudiney. (2019, July 25). *Penjelasan Sederhana tentang Apa Itu SVM? - Samsudiney - Medium*. Medium.

<https://medium.com/@samsudiney/penjelasan-sederhana-tentang-apa-itu-svm-149fec72bd02>

classification_report. (2025). Scikit-Learn. https://scikit-learn.org/1.5/modules/generated/sklearn.metrics.classification_report.html

