

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Sepak bola telah mengukuhkan posisinya sebagai salah satu cabang olahraga paling populer dan digemari di seluruh penjuru dunia. Di antara berbagai kompetisi liga profesional, English Premier League (EPL) secara konsisten menonjol sebagai salah satu liga yang paling banyak ditonton, memiliki daya tarik global, dan paling kompetitif. Popularitas EPL tidak hanya tecermin dari antusiasme miliaran penggemar, tetapi juga dari nilai ekonomi signifikan yang dihasilkannya, mencakup hak siar televisi yang bernilai miliaran dolar, penjualan merchandise, dan sponsor global. Menurut laporan Deloitte, pada musim 2022/2023, total pendapatan klub-klub EPL mencapai rekor tertinggi sebesar £6,1 miliar, meningkat 11% dari musim sebelumnya. Peningkatan ini terutama didorong oleh pendapatan dari hak siar internasional yang baru [1]. Selain itu, EPL disiarkan di lebih dari 180 negara, menjangkau audiens global yang luas [2].

Besarnya animo terhadap EPL berbanding lurus dengan tingginya minat untuk memprediksi berbagai aspek kompetisi, terutama hasil akhir klasemen. Prediksi ini menjadi topik hangat diskusi di kalangan penggemar, materi analisis mendalam bagi para pundit dan media olahraga, serta menjadi perhatian khusus dalam industri taruhan olahraga yang legal di berbagai negara [3]. Kemampuan untuk memprediksi peringkat tim di akhir musim menjadi tantangan menarik karena dinamika EPL yang sangat tinggi. Performa tim dipengaruhi oleh berbagai faktor kompleks dan seringkali sulit diprediksi, seperti kondisi fisik dan mental pemain, strategi pelatih, cedera pemain kunci, keputusan wasit, hingga performa tak terduga dari tim-tim non-unggulan. Contohnya, pada musim 2023/2024, persaingan untuk posisi empat besar berlangsung ketat hingga pekan terakhir, dengan selisih poin yang tipis antara beberapa tim [4]. Selain itu, kejutan dari tim-tim yang sebelumnya tidak diunggulkan menunjukkan betapa sulitnya memprediksi hasil akhir klasemen EPL.

Seiring dengan perkembangan zaman, upaya untuk melakukan prediksi dalam olahraga [5], termasuk sepak bola, telah berevolusi. Metode prediksi tradisional yang mengandalkan intuisi pakar atau analisis statistik sederhana seringkali memiliki keterbatasan dalam menangkap pola kompleks dari data historis

performa tim yang bersifat sekuensial dan dinamis[6]. Tingginya volume data yang tersedia saat ini, mulai dari statistik pertandingan, data pemain, hingga data historis pertemuan tim, membuka peluang untuk penerapan teknik analisis data yang lebih canggih. Salah satu pendekatan yang menunjukkan potensi besar dalam beberapa tahun terakhir adalah penggunaan machine learning (kecerdasan buatan). Menurut laporan dari Research and Markets, pasar analitik olahraga global diperkirakan akan tumbuh lebih dari \$15 miliar pada tahun 2030, didorong oleh kemampuan dan inovasi baru yang menyediakan data penting [7, 8, 9].

Dalam konteks prediksi data yang bersifat sekuensial atau time-series, seperti tren performa tim sepak bola sepanjang musim, salah satu algoritma machine learning yang sangat relevan adalah Long Short-Term Memory (LSTM). LSTM merupakan pengembangan dari Recurrent Neural Network (RNN) yang dirancang khusus untuk mengatasi masalah dalam mempelajari dependensi jangka panjang dalam data urutan waktu. Penelitian oleh Kharel et al. menunjukkan bahwa model LSTM efektif dalam menangkap pola dalam data time-series dan memberikan prediksi yang akurat dalam berbagai aplikasi, termasuk prediksi olahraga [10]. Kemampuan LSTM untuk "mengingat" informasi dari periode waktu yang lama dan melupakan informasi yang tidak relevan menjadikannya kandidat yang kuat untuk memodelkan fluktuasi performa tim dan memprediksi hasil akhir klasemen berdasarkan data historis pertandingan, statistik tim, dan variabel relevan lainnya [11, 12].

Meskipun beberapa penelitian telah mengeksplorasi penggunaan machine learning untuk prediksi hasil pertandingan sepak bola, penerapan LSTM secara spesifik untuk memprediksi peringkat akhir klasemen EPL secara komprehensif dengan memanfaatkan data historis terbaru masih memiliki ruang untuk dikembangkan lebih lanjut [3, 10]. Penelitian-penelitian sebelumnya mungkin memiliki fokus pada prediksi hasil pertandingan individual, menggunakan algoritma yang berbeda, atau cakupan data yang terbatas.

Oleh karena itu, penelitian ini diusulkan dengan tujuan untuk merancang dan membangun model prediksi hasil akhir klasemen English Premier League musim menggunakan algoritma Long Short-Term Memory (LSTM). Diharapkan model yang dikembangkan dapat memberikan kontribusi dalam memberikan perspektif prediksi yang lebih berbasis data dan akurat, serta menjadi referensi bagi penelitian selanjutnya di bidang analisis prediktif data olahraga.

1.2 Rumusan Masalah

1. Bagaimana melakukan prediksi hasil akhir klasemen EPL menggunakan algoritma LSTM?
2. Bagaimana kinerja model LSTM memprediksi hasil akhir klasemen EPL menggunakan metrik evaluasi *mean absolute error*, *mean squared error*, *Spearman correlation*, dan *Kendall Tau correlation*?

1.3 Batasan Permasalahan

1. Luas lingkup penelitian ini meliputi hasil akhir klasemen EPL 2024-2025 dan berfokus dengan penggunaan model LSTM.
2. Data yang digunakan untuk melatih model diambil dari Kaggle, situs resmi EPL, dan Fbref, kemudian dibatasi 25 musim sebelumnya.
3. Data yang digunakan hanya mencakup performa tim, tidak menggunakan statistik pemain.

1.4 Tujuan Penelitian

1. Menerapkan algoritma LSTM untuk memprediksi hasil akhir klasemen EPL.
2. Mengukur kinerja model LSTM memprediksi hasil akhir klasemen EPL menggunakan metrik *mean absolute error*, *mean squared error*, *Spearman correlation*, dan *Kendall Tau correlation*.

1.5 Manfaat Penelitian

Penelitian mengenai "Prediksi Hasil Akhir Klasemen English Premiere League Menggunakan Algoritma LSTM" memiliki beberapa manfaat sebagai berikut:

1. Memberikan wawasan tentang tren performa tim di EPL berdasarkan data historis yang bisa menjadi bahan diskusi komunitas sepak bola di Indonesia.
2. Jurnalis atau analis sepak bola di media olahraga dapat memanfaatkan penelitian ini sebagai bahan tambahan untuk membuat prediksi dan analisis yang lebih mendalam.

3. Hasil penelitian ini dapat menjadi referensi bagi mahasiswa lain yang ingin mendalami penggunaan machine learning, khususnya LSTM, dalam analisis prediktif berbasis data olahraga.

1.6 Sistematika Penulisan

Berisikan uraian singkat mengenai struktur isi penulisan laporan penelitian, dimulai dari Pendahuluan hingga Simpulan dan Saran.

Sistematika penulisan laporan adalah sebagai berikut:

- Bab 1 PENDAHULUAN
Bab ini berisi latar belakang masalah, rumusan masalah, batasan permasalahan, tujuan penelitian, manfaat penelitian, serta sistematika penulisan.
- Bab 2 LANDASAN TEORI
Bab ini membahas berbagai teori yang mendukung penelitian, seperti *deep learning*, LSTM, dan *english premiere league*.
- Bab 3 METODOLOGI PENELITIAN
Bagian ini menjelaskan metodologi yang digunakan dalam penelitian, mencakup tahapan studi literatur, pengumpulan dan analisis data, *data preprocessing*, arsitektur model, dan evaluasi.
- Bab 4 HASIL DAN DISKUSI
Bab ini hasil dari pemodelan prediksi yang telah dilakukan. Termasuk di dalamnya adalah hasil dari pemrosesan data, pemilihan fitur, *hyperparameter tuning*, serta evaluasi hasil prediksi dari model LSTM yang dibangun.
- Bab 5 KESIMPULAN DAN SARAN
Bab ini berisi simpulan dari hasil penelitian yang telah dilakukan serta saran untuk pengembangan atau penelitian selanjutnya.