

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terkait

Pertama, penelitian yang dilakukan oleh [23] bertujuan mengembangkan model peramalan debit *inflow* harian waduk menggunakan Jaringan Syaraf Tiruan (JST) yang dilatih dengan Algoritma Genetika (AG). Penelitian ini bertujuan meningkatkan pengelolaan waduk untuk fungsi seperti penyediaan energi, irigasi, dan pencegahan banjir, serta mengevaluasi pengaruh parameter genetika terhadap nilai *fitness* dalam pelatihan JST. Data curah hujan harian dan debit *inflow* digunakan sebagai input dan target dalam model JST. Proses pelatihan melibatkan pengkodean bobot dan bias ke dalam *kromosom*, di mana nilai *fitness* dihitung berdasarkan error hasil prediksi. Algoritma Genetika digunakan untuk mencari bobot optimal melalui seleksi, *crossover*, dan mutasi hingga beberapa generasi, menghasilkan model yang lebih akurat. Hasil menunjukkan pelatihan JST dengan AG menghasilkan nilai *fitness* optimal, dengan nilai terkecil 0,157 yang mencerminkan akurasi prediksi yang baik. Parameter genetika, seperti probabilitas *crossover* (0,9) dan jumlah generasi (50), berpengaruh signifikan terhadap nilai *fitness*, sementara probabilitas mutasi (0,5) dan ukuran populasi (100) hanya berdampak dalam rentang tertentu. Penelitian ini menegaskan efektivitas AG dalam meningkatkan akurasi prediksi debit inflow waduk dan memberikan wawasan penting untuk pengelolaan sumber daya air.

Kedua, penelitian yang dilakukan oleh [24] bertujuan untuk menganalisis dan mengoptimalkan *algoritma Neural Network* dengan fitur seleksi menggunakan algoritma genetika untuk memprediksi curah hujan, serta membandingkan kinerjanya dengan metode klasifikasi lain seperti *K-NN*, *Naive Bayes*, dan *Random Forest*. Penelitian ini juga bertujuan meningkatkan akurasi prediksi melalui pemilihan fitur yang relevan, sehingga dapat meningkatkan performa model dalam klasifikasi data curah hujan. Metode penelitian ini meliputi pembagian data curah hujan dari BMKG menjadi data pelatihan dan pengujian, diikuti dengan pra-

pemrosesan yang mencakup pengecekan *missing value*, penghapusan duplikat, dan normalisasi menggunakan *Z-Transformation*. Selanjutnya, dilakukan seleksi fitur menggunakan algoritma *Evolutionary*, diikuti dengan klasifikasi menggunakan algoritma Neural Network dan perbandingan dengan metode lain seperti K-NN, *Naive Bayes*, dan *Random Forest*. Validasi data dilakukan dengan metode 10-Fold Validation untuk mengevaluasi performa model berdasarkan *akurasi*, *presisi*, dan *recall*. Hasil penelitian menunjukkan bahwa algoritma Neural Network berbasis *Optimize Selection Evolutionary* mencapai akurasi tertinggi 81.13% setelah optimasi fitur, dan rata-rata akurasi 76.01% dengan *10-Fold Validation*. Pengujian dengan *Split Validation* rasio 0,6 menghasilkan akurasi 77.99%, dan rata-rata akurasi meningkat menjadi 83.61% setelah optimasi lebih lanjut. Kesimpulannya, *Neural Network* yang dioptimalkan dengan seleksi fitur efektif untuk memprediksi curah hujan.

Ketiga, penelitian yang dilakukan oleh [25] Penelitian ini bertujuan untuk mengembangkan model prediksi curah hujan di daerah Batu, Jawa Timur, dengan menggunakan metode *hybrid Fuzzy Inference System (FIS) Tsukamoto* dan algoritma genetika. Hal ini dilakukan untuk membantu petani apel dalam menentukan waktu pembungaan dan panen yang optimal. Dalam penelitian ini, metode yang digunakan adalah kombinasi antara *Fuzzy Inference System (FIS) Tsukamoto* dan algoritma genetika untuk memprediksi curah hujan di daerah Batu, Jawa Timur. Proses ini melibatkan optimasi batasan fungsi keanggotaan FIS Tsukamoto menggunakan algoritma genetika. Data historis curah hujan selama 10 tahun (2005-2014) dari empat lokasi pos hujan dianalisis untuk menghasilkan model prediksi. Hasilnya menunjukkan bahwa model hybrid ini memberikan nilai *error Root Mean Square Error (RMSE)* yang lebih kecil dibandingkan dengan FIS Tsukamoto tanpa optimasi, dengan nilai RMSE masing-masing untuk Junggo sebesar 6.485, Pujon 6.932, Tinjomulyo 5.969, dan Ngujung 5.498. Hal ini menunjukkan bahwa metode yang diusulkan efektif dalam meningkatkan akurasi prediksi curah hujan, yang sangat bermanfaat bagi para petani dalam menentukan waktu pembungaan dan panen. Penerapan metode hybrid ini efektif dalam memprediksi curah hujan dengan akurasi yang lebih tinggi. Optimasi menggunakan

algoritma genetika berhasil mengurangi error prediksi, yang bermanfaat bagi petani dalam merencanakan waktu pembungaan dan panen apel, sehingga hasil panen dapat dimaksimalkan.

2.2 Landasan Teori

2.2.1 Artificial Neural Network (ANN)

Artificial Neural Network (ANN) merupakan model komputasi yang terdiri dari sejumlah unit pemroses sederhana (neuron buatan) yang saling terhubung. Setiap neuron dalam ANN melakukan perhitungan terhadap input yang diterima dan menghasilkan output berdasarkan fungsi aktivasi[26]. Fungsi dasar dari *neuron* dalam ANN dapat dirumuskan sebagai berikut (1):

$$\begin{aligned} z^l &= w^l \cdot a^{l-1} + b^l \\ a^l &= f(z^l) \end{aligned} \quad (1)$$

Dimana w^l adalah bobot antara layer $l - 1$ ke layer l , b^l bias di layer l , a^{l-1} adalah *output* dari layer sebelumnya, f adalah fungsi aktivasi menggunakan ReLU dan Sigmoid.

2.2.2 Algoritma Genetika (AG)

Algoritma Genetika (AG) merupakan metode optimasi berbasis prinsip *evolusi biologis*, seperti seleksi alam, *crossover* (rekombinasi), dan mutasi, yang diterapkan untuk mencari kombinasi parameter terbaik dalam pelatihan model *Artificial Neural Network* (ANN). Dalam penelitian ini, AG digunakan untuk mengoptimalkan *hyperparameter* ANN, seperti jumlah *hidden layer*, jumlah neuron per layer, *dropout rate*, *learning rate*, dan *batch size*. [27]. Fungsi fitness dapat dirumuskan sebagai berikut, untuk kasus seperti prediksi banjir menggunakan MAE dan RMSE (2):

$$\text{Fitness} = \frac{1}{\text{MAE} + \text{RMSE} + \epsilon} \quad (2)$$

Dimana *Fitness* adalah Nilai kebugaran (*fitness*) dari sebuah individu (solusi ANN), MAE adalah *Mean Absolute Error* dari prediksi ANN, RMSE adalah *Root Mean Squared Error* dari prediksi ANN, ϵ adalah nilai konstan kecil 10^{-8} untuk mencegah pembagian dengan nol.

2.2.3 Metrik Evaluasi

Penggunaan metrik evaluasi memiliki peran penting dalam analisis data dan penerapan metode *Artificial Neural Network* (ANN) yang dioptimalkan dengan Algoritma Genetika (AG) untuk prediksi. Metrik ini digunakan untuk mengukur sejauh mana model yang dikembangkan mampu memprediksi data baru atau data yang belum pernah digunakan sebelumnya. Dengan metrik evaluasi, performa model dapat dinilai dalam mengenali pola dan hubungan yang terdapat dalam data. Selain itu, metrik evaluasi juga berguna untuk membandingkan berbagai konfigurasi model ANN yang dioptimalkan dengan AG sehingga memungkinkan pemilihan model yang paling akurat dan efisien[28].

Pada penelitian ini menggunakan metrik evaluasi yang digunakan untuk mengukur kinerja model antara lain sebagai berikut:

1. *Mean Absolute Error* (MAE) adalah rata-rata dari selisih absolut antara nilai aktual dan nilai prediksi. MAE memberikan gambaran tentang seberapa besar kesalahan prediksi secara keseluruhan. Rumus MAE adalah sebagai berikut (3):

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (3)$$

Dimana n adalah jumlah sampel dalam data, y^I adalah aktual dan \hat{y}^I adalah nilai prediksi.

2. *Root Mean Squared Error* (RMSE) adalah akar kuadrat dari MSE. RMSE memiliki interpretasi yang serupa dengan MSE, namun lebih mudah dipahami karena berada dalam satuan yang sama dengan data asli. Rumus RMSE adalah sebagai berikut (4):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (4)$$

Dimana n adalah jumlah data, i adalah urutan data pada database, y_i adalah aktual dan \hat{y}_i adalah nilai prediksi.

2.2.4 Prediksi Banjir

Prediksi banjir melibatkan analisis data historis seperti curah hujan, debit air, dan kondisi lingkungan lainnya. Dengan mengintegrasikan ANN dan AG, prediksi banjir dapat dilakukan dengan lebih akurat karena ANN mampu mengenali pola data yang kompleks, sementara AG membantu dalam mencari konfigurasi optimal model[29].

2.2.5 Machine Learning

Machine Learning merupakan cabang dari kecerdasan buatan (*Artificial Intelligence*) yang berfokus pada pengembangan sistem yang mampu belajar dan membuat keputusan berdasarkan data. Dalam konteks ini, komputer tidak diprogram secara eksplisit, melainkan dilatih untuk mengenali pola dari data historis yang tersedia. Salah satu pendekatan dalam *machine learning* adalah *supervised learning*, di mana model dilatih menggunakan data yang sudah dilabeli. Metode seperti *Artificial Neural Network* (ANN) termasuk dalam kategori ini dan banyak digunakan untuk tugas prediksi, termasuk prediksi kejadian bencana seperti banjir. Untuk meningkatkan akurasi prediksi, ANN dapat dioptimalkan dengan Algoritma Genetika (AG), yang berfungsi untuk mencari parameter terbaik dalam proses pelatihan model. Pendekatan ini memungkinkan pemrosesan data dalam skala besar dan mendukung pengambilan keputusan berbasis data secara lebih efektif dalam mitigasi risiko bencana[30].

2.2.6 Python

Python merupakan bahasa pemrograman yang banyak digunakan dalam pengembangan model prediktif, analisis data, dan kecerdasan buatan karena sintaksisnya yang sederhana serta *fleksibel*. *Python* memiliki berbagai pustaka (*library*) yang mendukung implementasi metode *machine learning* dan optimasi, seperti *Pandas* untuk manipulasi data, *NumPy* untuk komputasi numerik, *Matplotlib* dan *Seaborn* untuk visualisasi data, serta *TensorFlow* dan *Keras* untuk membangun model *Artificial Neural Network*. Selain itu, pustaka seperti *DEAP* atau *PyGAD* juga dapat dimanfaatkan untuk penerapan Algoritma Genetika dalam proses optimasi. Dalam penelitian ini, *Python* digunakan sebagai alat bantu utama untuk mengolah data historis banjir, membangun arsitektur ANN, mengintegrasikan

optimasi menggunakan Algoritma Genetika, serta mengevaluasi kinerja model secara kuantitatif[31].

2.2.7 Visual Code

Visual Studio Code (VS Code) adalah editor kode sumber yang dikembangkan oleh Microsoft dan banyak digunakan dalam pengembangan perangkat lunak karena sifatnya yang ringan, gratis, dan mendukung lintas platform. VS Code memiliki berbagai fitur seperti *syntax highlighting*, *debugging*, *terminal terintegrasi*, serta dukungan ekstensi yang luas untuk berbagai bahasa pemrograman, termasuk *Python*. Dalam penelitian ini, VS Code digunakan sebagai lingkungan kerja utama untuk menulis, menjalankan, dan menguji kode *Python* dalam membangun model *Artificial Neural Network (ANN)* yang dioptimalkan menggunakan Algoritma Genetika (GA). Kemampuannya untuk terintegrasi dengan pustaka seperti *Pandas*, *NumPy*, *TensorFlow/Keras*, dan *PyGAD* menjadikan VS Code sebagai alat yang efektif dalam proses pengolahan data historis, pelatihan model prediktif, serta evaluasi performa prediksi bencana banjir[32].