

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi yang pesat, mendorong pemanfaatannya dalam berbagai bidang kehidupan, termasuk dalam dunia survei. Penggunaan teknologi informasi dalam survei telah menghasilkan kemajuan signifikan dalam pengumpulan, analisis, dan interpretasi data hasil survei. Salah satu aspek penting dalam survei adalah penanganan data yang tidak wajar atau anomali, serta keberadaan data yang hilang atau kosong. Kemajuan teknologi dalam bidang Jaringan Syaraf Tiruan memberikan kemudahan dalam menangani hal ini.

Survei merupakan salah satu instrumen yang digunakan dalam pengumpulan data. Data hasil survei diolah dan dianalisis untuk mendapatkan karakteristik dari suatu populasi. Kualitas data yang dihasilkan dari kegiatan survei, sering dinilai dari besarnya nilai bias yang terbentuk. Menurut Felderel, *et al.* (2019), nilai bias dapat diukur dengan akurat apabila peneliti memahami sumber *error* yang menyebabkan bias, yaitu ketidaktersediaan data dan anomali data atau pengukuran yang salah.

Menurut Carpenter dan Smuk (2021), ketidaktersediaan data menjadi hal yang tidak dapat dihindari dan sering terjadi dalam penelitian. Analisis terhadap data bisa menjadi rumit dan sering menimbulkan keraguan akibat adanya ketidaktersediaan data. Responden yang menolak untuk didata (*unit nonresponse*) atau responden tidak memberikan seluruh data yang ditanyakan (*item nonresponse*) menjadi salah satu penyebabnya (Kalpourtzzi, *et al.*, 2023).

Anomali data dalam konteks survei merujuk pada nilai-nilai yang tidak wajar atau tidak sesuai dengan pola yang diharapkan dalam kumpulan data. Anomali data dapat disebabkan oleh berbagai faktor, seperti kesalahan pengisian oleh

responden, kesalahan pengumpulan data, atau adanya karakteristik khusus dari populasi yang sedang disurvei. Menurut Gajendra, *et al.* (2020), deteksi anomali data menjadi hal yang penting dilakukan karena dapat mempengaruhi hasil suatu analisis statistik. Anomali data bisa menyebabkan kesalahan dalam membuat model data, estimasi parameter, dan hasil peramalan yang tidak sesuai standar.

Salah satu metode yang dapat digunakan dalam melakukan prediksi anomali dan ketidaktersediaan data adalah Metode *Backpropagation*. Metode *Backpropagation* merupakan bagian dari Jaringan Syaraf Tiruan (*artificial neural networks*) yang digunakan untuk mempelajari pola dan melakukan prediksi berdasarkan data yang ada. Menurut Chauvin dan Rumelhart (1995), metode ini bekerja dengan menentukan fungsi *error* dan menggunakan *gradient descent* untuk melakukan perbaikan *weight*. Perbaikan *weight* dilakukan agar *network* bisa menghasilkan performa yang lebih baik dalam mendekati nilai output.

Penelitian yang dilakukan oleh Gajendra, *et al.* (2020) menggunakan Metode *Backpropagation* untuk menentukan anomali data pada data perdagangan aluminium tahun 2006-2015 di India. Penelitian ini membuktikan bahwa Metode *Backpropagation* terbukti memiliki nilai efisiensi dalam penentuan anomali data yang lebih baik dibanding Metode *Bayesian approach to detect abnormal regions* (BARD), *Package R Otsad*, *Chen and Liu algorithm*, dan *MHD Detection*. Keunggulan ini dapat diperoleh karena metode ini mampu menentukan lokasi data anomali secara tepat. Metode lain hanya bisa menentukan lokasi data anomali secara rentang nilai.

Sebayang dan Yuniarto (2017) melakukan penelitian menggunakan Metode *Backpropagation* yang ditingkatkan dengan Algoritma Genetika. Optimasi dilakukan untuk mengatasi beberapa kelemahan Metode *Backpropagation*, seperti membutuhkan ketepatan dalam penentuan topologi jaringan dan parameter pembelajaran. Jika jaringan telah dilatih dan akurasi yang dihasilkan tidak layak, peneliti harus mengulang kembali penghitungan dengan nilai parameter yang berbeda. Parameter yang diganti berupa jumlah *hidden layer*, nilai bobot awal, dan nilai *learning rate*. Penelitian ini menggunakan data simulasi yang dibangkitkan mengikuti sebaran distribusi normal. Penelitian membuktikan bahwa dengan tingkat kepercayaan 95%, metode yang diteliti memiliki nilai akurasi yang lebih baik dibanding penghitungan menggunakan Metode Regresi Linear Berganda.

Mustikaningrum dan Wardoyo (2020) meneliti pengelompokan penyakit kanker yang menyerang sel darah putih menggunakan Metode *Backpropagation* yang ditingkatkan dengan Algoritma Genetika. Algoritma Genetika digunakan dalam pemilihan arsitektur, *learning rate*, dan momentum yang optimal. Hasil penelitian ini menunjukkan bahwa pelatihan dan pengujian dengan optimasi Algoritma Genetika terhadap parameter Metode *Backpropagation* menghasilkan rata-rata akurasi validasi sebesar 94,3%. Metode tersebut terbukti lebih baik dibandingkan dengan Metode *Backpropagation* biasa yang menghasilkan rata-rata akurasi validasi sebesar 88,22%.

Zhang dan Qu (2021) melakukan prediksi menggunakan data arus lalu lintas di distrik Xian, China. Penelitian menggunakan Metode *Backpropagation* yang ditingkatkan dengan Algoritma Genetika. Peneliti menggunakan Algoritma Genetika untuk memilih variabel terbaik yang sangat mempengaruhi data hasil. Variabel tersebut kemudian digunakan sebagai *input* dalam melakukan prediksi pada Metode *Backpropagation*. Penelitian ini membuktikan bahwa Metode *Backpropagation* yang ditingkatkan dengan Algoritma Genetika memiliki akurasi penghitungan yang lebih baik apabila dibandingkan dengan penghitungan tanpa optimasi Algoritma Genetika.

Penelitian yang dilakukan oleh Chen, *et al.* (2019) menggunakan Algoritma Genetika untuk meningkatkan efisiensi Metode *Backpropagation*. Penelitian ini bertujuan untuk memprediksi ketinggian air menggunakan data kondisi perairan pada semua stasiun Meteorologi di Han River Basin, distrik Hubei, China. Algoritma Genetika digunakan untuk menentukan parameter bobot awal dan *bias* terbaik yang akan digunakan dalam penghitungan Metode *Backpropagation*. Hasil penelitian ini membuktikan bahwa nilai efisiensi Metode *Backpropagation* dapat lebih baik apabila ditingkatkan menggunakan Algoritma Genetika..

Salah satu bahasa pemrograman yang digunakan untuk melakukan penghitungan Metode *Backpropagation* adalah bahasa R. Menurut Wallace, *et al.* (2012), ada banyak manfaat dalam menggunakan R sebagai mesin komputasi statistik. Pengguna dapat memanfaatkan banyak metode statistik yang telah ditulis dalam bahasa ini. Pengguna tidak perlu melakukan *coding* dari awal. R adalah bahasa pemrograman *open-source*, yang berarti *codingan* tersedia untuk umum dan dapat diakses dan dimodifikasi oleh siapa saja. Pengguna dapat berkontribusi pada pengembangan R dengan menambahkan fungsi-fungsi baru atau memperbaiki yang

sudah ada. R dapat digunakan pada berbagai sistem operasi dan mudah diterapkan pada aplikasi berbasis *desktop* maupun *website*.

Keberadaan komunitas R yang besar dan aktif membuat R menjadi pilihan yang kuat dalam analisis statistik dan penelitian ilmiah. Komunitas ini menyediakan dukungan, sumber daya, dan forum diskusi yang dapat digunakan pengguna R untuk berbagi pengetahuan, mengajukan pertanyaan, dan memperoleh pemecahan masalah. Interaksi dengan komunitas R ini membantu pengguna untuk terus belajar dan mengembangkan keterampilan dalam menggunakan R sebagai alat statistik.

Berdasarkan uraian di atas, penulis tertarik melakukan penelitian dalam bentuk tesis dengan judul “Peningkatan Efektivitas Metode *Backpropagation* Menggunakan Algoritma Genetika untuk Melakukan Prediksi Anomali dan Ketidakterersediaan Data Hasil Survei (Studi Kasus di BPS Kabupaten Solok). Penulis menggunakan data hasil Survei Sosial Ekonomi Nasional (Susenas) yang dilakukan oleh Badan Pusat Statistik (BPS) Kabupaten Solok sebagai objek penelitian. Menurut BPS (2022b), kegiatan survei Susenas bertujuan untuk mengumpulkan data dalam membantu penyusunan kebijakan publik dan perencanaan pembangunan. Penelitian ini diharapkan dapat membantu Tim Kerja Pengolahan Data BPS dalam melakukan pengecekan anomali data dan *professional adjustment* bagi data yang tidak tersedia.

1.2 Perumusan Masalah

Berdasarkan uraian latar belakang di atas, penulis merumuskan beberapa permasalahan sebagai berikut:

1. Bagaimana memahami penerapan Metode *Backpropagation* dalam melakukan prediksi anomali dan ketidakterersediaan data?
2. Bagaimana menganalisa kinerja Metode *Backpropagation* yang ditingkatkan dengan Algoritma Genetika dalam melakukan prediksi anomali dan ketidakterersediaan data hasil survei?
3. Bagaimana membangun aplikasi dengan bahasa pemrograman R untuk melakukan prediksi anomali dan ketidakterersediaan data?
4. Bagaimana menguji keakuratan dan ketepatan hasil aplikasi dalam melakukan prediksi menggunakan Metode *Backpropagation* yang ditingkatkan dengan Algoritma Genetika?

1.3 Batasan Masalah

Penulis membatasi ruang lingkup objek penelitian agar pembahasan penelitian tidak menyimpang dari rumusan masalah di atas. Penelitian ini hanya berfokus pada:

1. Data hasil Susenas BPS Kabupaten Solok sebagai objek penelitian.
2. Aplikasi yang digunakan menggunakan bahasa pemrograman R.
3. Tempat penelitian dilakukan di BPS Kabupaten Solok.

1.4 Tujuan Penelitian

Tujuan penelitian adalah rumusan yang menjelaskan hasil yang ingin dicapai penulis melalui pelaksanaan penelitian. Tujuan penelitian memberikan arah dan fokus bagi penulis serta menentukan hasil yang diharapkan. Adapun tujuan yang ingin dicapai adalah:

1. Memahami penerapan Metode *Backpropagation* dalam melakukan prediksi anomali dan ketidaktersediaan data.
2. Menganalisa kinerja Metode *Backpropagation* yang ditingkatkan dengan Algoritma Genetika dalam melakukan prediksi anomali dan ketidaktersediaan data hasil survei.
3. Membangun aplikasi dengan bahasa pemrograman R untuk melakukan prediksi anomali dan ketidaktersediaan data.
4. Menguji keakuratan dan ketepatan hasil aplikasi dalam melakukan prediksi menggunakan Metode *Backpropagation* yang ditingkatkan dengan Algoritma Genetika.

1.5 Manfaat Penelitian

Manfaat penelitian adalah hasil positif yang diperoleh dari suatu penelitian yang memberikan kontribusi, nilai tambah, atau solusi bagi perkembangan ilmiah, praktis, kebijakan, atau pengambilan keputusan. Adapun manfaat penelitian yang ingin dicapai oleh penulis adalah sebagai berikut:

1. Memberikan referensi bagi peneliti lain dalam melakukan peningkatan performa Metode *Backpropagation* menggunakan Algoritma Genetika.
2. Membantu Tim Kerja Pengolahan Data BPS dengan menyediakan data prediksi dalam melakukan pengecekan anomali data dan *professional adjustment* bagi data yang tidak tersedia.

1.6 Sistematika Penulisan

Sistematika penulisan menjadi panduan urutan penelitian yang dilakukan penulis. Sistematika penulisan penelitian ini adalah sebagai berikut:

BAB I. PENDAHULUAN

Bab ini berisikan tentang latar belakang penelitian, perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB II. LANDASAN TEORI

Bab ini menjelaskan Metode *Backpropagation* dan Algoritma Genetika serta teori-teori dan penerapan metode yang berhubungan dalam penelitian yang dilakukan.

BAB III. METODOLOGI PENELITIAN

Bab ini berisikan kerangka kerja, mendefinisikan ruang lingkup, analisa masalah, menentukan tujuan, mempelajari literatur, mengumpulkan data dan informasi, menganalisa dan menerapkan Metode *Backpropagation* dan Algoritma Genetika, mengimplementasikan metode, menguji data dan menarik kesimpulan.

BAB IV. ANALISA DAN PERANCANGAN

Bab ini berisikan analisa dan penerapan metode yang meliputi gambaran umum penelitian, data yang digunakan, tahap penelitian, spesifikasi alat, dan gambaran perancangan.

BAB V. IMPLEMENTASI DAN HASIL

Bab ini menjelaskan tentang implementasi dan penerapan Metode *Backpropagation* dan Algoritma Genetika dalam melakukan prediksi anomali dan ketidakterersediaan data hasil survei.

BAB VI. KESIMPULAN DAN SARAN

Bab ini berisikan jawaban singkat dari rumusan masalah atau kesimpulan serta saran dalam pengembangan dan penyempurnaan penelitian di masa yang akan datang.