

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pengambilan keputusan dari suatu permasalahan merupakan hal yang tidak dapat dipisahkan dalam kehidupan. Proses pengambilan keputusan pada suatu permasalahan bertujuan untuk menentukan solusi permasalahan tersebut. Penentuan solusi dan pengambilan suatu kebijakan dapat diketahui berdasarkan karakteristik permasalahannya. Teknik segmentasi merupakan pendekatan yang dapat diterapkan untuk mengetahui karakteristik dari suatu permasalahan tersebut. Segmentasi akan membagi satu kelompok populasi menjadi kelompok-kelompok yang berbeda karakteristiknya antara satu dengan yang lainnya. Menurut Kasali (1998), terdapat dua jenis metode yang dapat digunakan untuk segmentasi, yaitu metode interdependensi dan metode dependensi. Proses pembentukan segmentasi dengan metode interdependensi hanya dari variabel-variabel independen yang tidak dikaitkan dengan variabel dependennya. Sedangkan, pada metode dependensi segmen-segmen dihasilkan karena terdapat hubungan antara variabel dependen dengan variabel-variabel independennya. Salah satu teknik segmentasi metode dependensi adalah *Chi-Squared Automatic Interaction Detection* (Gallagher, Monroe, & Fish, 2000)

Metode CHAID merupakan suatu teknik iteratif yang menguji satu persatu variabel independen yang digunakan dalam klasifikasi, dan menyusunnya berdasarkan pada tingkat signifikansi statistik *chi-square* terhadap variabel dependennya (Gallagher, Monroe, & Fish, 2000). Dalam metode ini pembentukan segmentasi dilakukan setelah terbentuk diagram pohon, dimana percabangan variabel-variabel independen pada diagram pohon dilakukan menggunakan uji *chi-square* (Gallagher, Monroe, & Fish, 2000).

Penerapan metode *Chi-Squared Automatic Interaction Detection* (CHAID) digunakan data kategorik. Pada data kategorik, sangat rentan terjadi

ketidakseimbangan jumlah data antara kelas mayor dengan kelas minor. Kelas mayor adalah kelas yang memiliki jumlah data lebih besar daripada kelas minor. Ketidakseimbangan data terjadi ketika kategori klasifikasi tidak direpresentasikan secara merata (Chawla, Bowyer, Hall, & Kegelmeyer, 2002). Barro, Sulvianti (2013) menyatakan bahwa data yang tidak seimbang dapat meningkatkan kesalahan klasifikasi dalam kelas minor dan model yang terbentuk lebih mengarah pada kelas mayoritas.

Pemecahan masalah untuk mengatasi data tidak seimbang terdapat beberapa pendekatan, salah satunya adalah *resampling*. Ada dua jenis *resampling*, yaitu *undersampling* dan *oversampling*. *Undersampling* adalah metode yang bekerja dengan cara mengurangi kelas mayoritas sampai memiliki jumlah data yang sama dengan kelas minoritasnya. Sedangkan *oversampling* dilakukan pada kelas minor dengan mereplikasi sampai memiliki proporsi yang seimbang. *Undersampling* sangat jarang dilakukan karena dapat mengurangi atau mengambil informasi penting dari kelas mayoritas. Namun, *oversampling* juga dapat meningkatkan kemungkinan terjadinya *overfitting*, karena menduplikasi beberapa *instance* yang sama sehingga tidak memiliki variasi data. Karena kelemahan *undersampling* dan *oversampling*, Chawla, dkk (2002) mengusulkan metode *Synthetic Minority Oversampling Technique* (SMOTE) yang merupakan pendekatan untuk mengatasi ketidakseimbangan data sebagai perbaikan dari metode *oversampling*.

Penerapan metode SMOTE digunakan untuk data berskala numerik. Namun, pada realitanya sering kali data yang didapat berupa data kategorik, sehingga Chawla melakukan pengembangan dari metode SMOTE yang digunakan untuk data berskala berbentuk kategorik. Metode ini bekerja dengan menggunakan *k-nearest neighbors* untuk membuat *synthetic data*. Pada metode SMOTE untuk data kategorik, *nearest neighbors* dihitung menggunakan versi modifikasi dari *Value Difference Metric* (VDM) (Chawla, Bowyer, Hall, & Kegelmeyer, 2002).

Ketidakseimbangan data biasanya muncul pada kasus-kasus di bidang kesehatan atau suatu kasus penyakit (Arifiyanti, 2020). Hal ini muncul karena

Ivo Oktaviano, 2021

PENGLASIFIKASIAN MENGGUNAKAN CHI-SQUARE AUTOMATIC INTERACTION DETECTION DENGAN PENDEKATAN SYNTHETIC MINORITY OVERSAMPLING TECHNIQUE (Studi Kasus Bayi Berat Lahir Rendah (BBLR) di Jawa Barat)

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

kasus terkena penyakit yang ingin diprediksi jumlah kemunculannya lebih sedikit jika dibandingkan dengan kasus yang tidak terkena penyakit. Dalam penelitian ini, akan dilakukan analisis CHAID dengan pendekatan SMOTE dengan mengambil studi kasus yaitu Bayi Berat Lahir Rendah (BBLR) di Jawa Barat.

Provinsi Jawa Barat merupakan salah satu Provinsi di Indonesia, tentunya Indonesia memiliki tujuan-tujuan yang hendak dicapai. Upaya untuk mencapai tujuan tersebut dapat dilakukan melalui pembangunan di berbagai bidang. Salah satu bentuk program dan rencana pembangunan yang dilaksanakan oleh pemerintah adalah Rencana Pembangunan Jangka Menengah Nasional (RPJMN). Sebagai arah pembangunan nasional, RPJMN terintegrasi dengan *Sustainable Development Goals* (SDGs) yang memuat tujuan dan sasaran pembangunan. Salah satu tantangan yang dihadapi oleh pemerintah dalam proses pelaksanaan RPJMN adalah meningkatkan sumber daya manusia (SDM). Peningkatan sumber derajat kesehatan merupakan salah satu upaya yang dilakukan pemerintah untuk meningkatkan sumber daya manusia (BAPPENAS, 2015).

Derajat kesehatan terdapat tiga indikator yaitu mortalitas, morbiditas, dan status gizi (Departemen Kesehatan RI, 2003). Mortalitas sebagai indikator derajat kesehatan ini salah satunya terdiri dari Angka Kematian Bayi (AKB) dan Angka Kematian Neonatal (AKN). AKB dan AKN merupakan tolak ukur dari status kesehatan anak/bayi dan penilaian keberhasilan dalam pembangunan nasional. Salah satu indikator dalam kesehatan bayi baru lahir adalah berat badan. Bayi Berat Lahir Rendah (BBLR) merupakan masalah yang perlu diperhatikan karena bayi dengan berat lahir rendah lebih besar risikonya untuk mengalami komplikasi saat lahir bahkan memiliki risiko lebih tinggi terhadap kematian usia dini. (Kemenkes, 2016)

Menurut Survey Demografi dan Kesehatan Indonesia (SDKI) 2017 diperoleh informasi bahwa angka BBLR di Indonesia adalah sekitar 7%, artinya terdapat 7 bayi terlahir dalam kondisi BBLR dari 100 kelahiran hidup. Berdasarkan hasil analisa lanjut yang dilakukan SDKI 2017 periode 5 tahun

Ivo Oktaviano, 2021

PENGLASIFIKASIAN MENGGUNAKAN CHI-SQUARE AUTOMATIC INTERACTION DETECTION DENGAN PENDEKATAN SYNTHETIC MINORITY OVERSAMPLING TECHNIQUE (Studi Kasus Bayi Berat Lahir Rendah (BBLR) di Jawa Barat)

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

yaitu 2012 sampai dengan 2017, diperoleh informasi bahwa Angka Kematian Bayi (AKB) di Indonesia mencapai 24 kematian per 1000 kelahiran hidup, hal ini berarti 1 dari 42 anak meninggal sebelum ulang tahun pertamanya. Sedangkan Angka Kematian Neonatal (AKN) di Indonesia adalah 15 kematian per 1000 kelahiran hidup, atau artinya 1 dari 67 anak meninggal dalam bulan pertama kehidupannya. Berdasarkan profil kesehatan Indonesia tahun 2017, Provinsi Jawa Barat merupakan salah satu provinsi yang mempunyai persentase BBLR sebesar 7,1% atau sebanyak 71 kejadian BBLR per 1000 kelahiran hidup. Angka ini menunjukkan bahwa BBLR belum mencapai target SDGs yaitu sebesar 5-6% untuk periode 2012-2017, sehingga perlu upaya untuk menurunkan angka BBLR.

Sebagaimana paparan di atas diketahui bahwa kejadian BBLR 2017 di Jawa Barat masih menjadi permasalahan kesehatan. Angka BBLR 2017 di Indonesia, khususnya di Provinsi Jawa Barat masih belum memenuhi target dari SDGs, sehingga perlu adanya penelitian yang bertujuan mengurangi angka BBLR agar mencapai tujuan dari keberhasilan pembangunan nasional. Berdasarkan uraian sebelumnya, maka peneliti akan melakukan penelitian yang berjudul “Pengklasifikasian Menggunakan *Chi-Square Automatic Interaction Detection* dengan Pendekatan *Synthetic Minority Oversampling Technique* (Studi Kasus Bayi Berat Lahir Rendah (BBLR) di Jawa Barat)”.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang telah dijelaskan, maka rumusan masalah dalam penelitian ini, yaitu:

1. Bagaimana implementasi dari pengklasifikasian menggunakan metode *Chi Square Automatic Interaction Detection* dan *Chi Square Automatic Interaction Detection* dengan pendekatan *Synthetic Minority Oversampling Technique* pada kasus BBLR di Jawa Barat?
2. Bagaimana hasil dari perbandingan metode *Chi Square Automatic Interaction Detection* dan *Chi Square Automatic Interaction Detection*

Ivo Oktaviano, 2021

PENGLASIFIKASIAN MENGGUNAKAN CHI-SQUARE AUTOMATIC INTERACTION DETECTION DENGAN PENDEKATAN SYNTHETIC MINORITY OVERSAMPLING TECHNIQUE (Studi Kasus Bayi Berat Lahir Rendah (BBLR) di Jawa Barat)

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

dengan pendekatan *Synthetic Minority Oversampling Technique* pada kasus BBLR di Jawa Barat?

1.3 Batasan Masalah

Berdasarkan latar belakang di atas, maka batasan masalah penelitian ini adalah responden dan variabel independen. Responden dalam penelitian ini adalah ibu yang memiliki anak di Provinsi Jawa Barat yang tercatat saat wawancara SDKI 2017. Variabel independen yang diteliti meliputi umur ibu, pendidikan ibu, tingkat ekonomi, paritas ibu, tempat tinggal, status bekerja ibu, serta kebiasaan merokok ibu. Variabel independen yang tidak diteliti merupakan variabel yang tidak termasuk dalam data SDKI 2017, yaitu meliputi kadar Hb, penyakit saat mengandung, paparan radiasi, faktor janin, dan faktor placenta.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah penelitian di atas, maka tujuan penelitian ini, yaitu sebagai berikut:

1. Mengimplementasikan pengklasifikasian menggunakan metode *Chi Square Automatic Interaction Detection* dan *Chi Square Automatic Interaction Detection* dengan pendekatan *Synthetic Minority Oversampling Technique* pada kasus BBLR di Jawa Barat
2. Membandingkan hasil *Chi Square Automatic Interaction Detection* pada kasus BBLR di Jawa Barat dan *Chi Square Automatic Interaction Detection* dengan pendekatan *Synthetic Minority Oversampling Technique* pada kasus BBLR di Jawa Barat.

1.5 Manfaat Penelitian

Manfaat yang diharapkan dalam penelitian ini adalah sebagai berikut:

1.5.1 Manfaat Teoritis

Secara teoritis, penulisan skripsi ini diharapkan dapat memberikan tambahan informasi tentang penerapan metode *Chi Square Automatic*

Ivo Oktaviano, 2021

PENGLASIFIKASIAN MENGGUNAKAN CHI-SQUARE AUTOMATIC INTERACTION DETECTION DENGAN PENDEKATAN SYNTHETIC MINORITY OVERSAMPLING TECHNIQUE (Studi Kasus Bayi Berat Lahir Rendah (BBLR) di Jawa Barat)

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Interaction Detection dengan pendekatan *Synthetic Minority Oversampling Technique* pada kasus BBLR di Jawa Barat.

1.5.2 Manfaat Praktis

Bagi pembaca khususnya pemerintah Provinsi Jawa Barat, diharapkan penulisan skripsi ini dapat menambah informasi serta wawasan tentang seberapa pentingnya BBLR harus ditangani di Provinsi Jawa Barat.