

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam perkembangan zaman, semakin hari semakin banyak inovasi dalam bidang teknologi tak terkecuali dalam bidang komputerisasi. Dengan semakin canggihnya perkembangan teknologi saat ini maka dapat di terapkan ke dalam berbagai bidang, termasuk salah satunya pada bidang analisis data. Dalam bidang ilmu statistik sangat dibutuhkan peran teknologi untuk memudahkan dan membantu peneliti dalam menyelesaikan suatu permasalahan. Ditunjang dengan kecepatan dan keakuratan teknologi saat ini, banyak metode statistik yang dapat dikembangkan untuk mendapatkan hasil analisis terbaik dari tiap pengolahan data. Dengan adanya bantuan dari komputer, penelitian dapat dengan mudah mengolah data besar atau rumit dengan tingkat kesalahan yang dapat diminimalisir. Oleh karena itu komputerisasi sangatlah memudahkan penelitian dalam kasus penghitungan maupun replikasi rumit.

Metode statistika mengalami perkembangan yang pesat diiringi semakin kompleksnya masalah dalam kehidupan yang harus segera di carikan solusi penyelesaiannya, dengan cara melakukan penelitian. Statistika ikut berperan penting dalam pemecahan masalah seperti halnya penggunaan analisis regresi. Analisis regresi merupakan salah satu metode statistika yang digunakan untuk mengetahui suatu hubungan sebab akibat (kausalitas) antara satu variabel dependen

dengan satu atau lebih variabel independen. Variabel dependen dapat berupa data kategorik lebih dari satu, maka digunakan analisis regresi logistik. Jika variabel dependen hanya memiliki 2 kategori/kemungkinan (*dichotomous*) contoh, variabel ujian masuk perguruan tinggi (Y) mempunyai 2 kategori yaitu lulus dan tidak lulus, variabel penggunaan alat kontrasepsi dengan kategori menggunakan dan tidak menggunakan, maka metode yang digunakan adalah analisis regresi logistik biner. Metode yang digunakan untuk menyelesaikan masalah regresi logistik terbaik yakni dengan mencari model yang dapat menggambarkan masalah kausalitas, adalah metode *Maximum Likelihood Estimation* (MLE) dimana proses penaksiran parameter didahului oleh pembentukan fungsi *likelihood*.

Bootstrap adalah metode dan teknik *resampling* nonparametrik yang berbasis komputer dan sangat baik untuk dipergunakan pada masalah keakurasian yang didasarkan pada simulasi data untuk keperluan inferensi statistik (Sungkono, 2013). *Bootstrap* diperkenalkan pertama kali oleh Efron tahun 1979. *Bootstrapping* bertujuan untuk menentukan estimasi yang kuat dari standar error dan interval kepercayaan untuk mengestimasi proporsi, rerata, median, odds ratio, koefisien korelasi atau koefisien regresi (Widhiarso, 2012). Jika dalam penelitian kita hanya memiliki sampel kecil maka *Bootstrap* dapat mengatasinya, selain itu juga dapat mengatasi permasalahan data yang menyimpang dari asumsinya maupun data yang tidak memiliki atau bebas asumsi dalam distribusinya (Karomah & Hendikawati, 2013). Banyak juga kasus pada penelitian regresi logistik yang lampau, data yang kelasnya tidak seimbang tetapi dari pihak peneliti mengabaikan adanya

ketidakseimbangan kelas tersebut, sehingga pada hasil prediksi cenderung condong pada kelas mayoritas. Untuk mengatasi masalah tersebut, metode *resampling* merupakan suatu metode dapat digunakan untuk memecahkan masalah ketidakseimbangan kelas (Rianto, 2015).

Metode *Bootstrap* dilakukan dengan mengambil sampel dari sampel asli dengan ukuran sama dengan ukuran sampel asli dan dilakukan dengan pengembalian. Dinotasikan dengan $X_1^*, X_2^*, \dots, X_B^*$, masing-masing berisi nilai data yang telah disampling secara random dengan pengembalian dari sampel x . Kemudian evaluasi hasil *Bootstrap* yang diperoleh untuk masing-masing sampel *Bootstrap*. Lalu estimasi *standard error* untuk sampel *Bootstrap*, langkah tersebut dilakukan dengan pengulangan beberapa kali hingga di dapatkan hasil terbaik. Keuntungan metode *resampling Bootstrap* adalah nilai *standard error* akan mengalami penurunan secara signifikan jika diaplikasikan pada data atau jumlah yang kecil.

Menurut Ridha Ramandhani, dalam kasus ketepatan klasifikasi kesejahteraan rumah tangga di kota Pati, setelah melakukan analisis dengan metode *Bootstrap Aggregating (Bagging)* regresi logistik biner, diperoleh tiga variabel prediktor yang berpengaruh signifikan terhadap kesejahteraan rumah tangga di kota Pati yaitu jenis kelamin kepala keluarga, jumlah anggota ruma tangga, dan penguasaan telepon seluler dengan tingkat akurasi sebesar 79,87%. Hasil analisis regresi logistik biner

dengan replikasi *Bootstrap* sebanyak 50, 60, 70, 80, 90, 100, 150, 626, dan 1000 kali menunjukkan bahwa terdapat konsistensi pada setiap pengulangan.

Menurut penelitian Ahmad Reza Aditya berjudul ketepatan klasifikasi pemilihan metode kontrasepsi di kota Semarang menggunakan *Bootstrap Aggregating Regresi Logistik Multinomial* mendapatkan hasil akhir bahwa model yang terbentuk terhadap metode kontrasepsi (Y) yaitu jumlah anak, pendidikan suami, pendidikan istri, dan tingkat kesejahteraan keluarga. Perbandingan ketepatan klasifikasi model regresi logistik multinomial dengan *Bootstrap Aggregating regresi logistik Multinomial* diperoleh bahwa model terbaik adalah model dari *Bootstrap Aggregating regresi logistik multinomial* dengan replikasi sebanyak 50 kali dengan ketepatan klasifikasi terbesar adalah 51%.

Kecelakaan lalu lintas merupakan momok mengerikan yang terjadi di banyak negara, terlebih untuk negara-negara berkembang ketika masalah transportasi seperti benang kusut. Angka kematian global di sebabkan kecelakaan lalu lintas saat ini tercatat mencapai angka 1,24 juta per tahun. Diperkirakan, angka tersebut akan meningkat hingga tiga kali lipat menjadi 3,6 juta per tahun pada 2030. Di negara berkembang kecelakaan lalu lintas termasuk lima besar penyebab utama kematian di dunia. Melampaui HIV/AIDS, malaria, TBC, dan penyakit pembunuh lainnya. Data yang dikeluarkan *World Health Organization* (WHO) menunjukkan, India menempati urutan pertama negara dengan jumlah kematian terbanyak akibat kecelakaan lalu lintas sementara Indonesia menempati urutan kelima. Namun Indonesia justru menempati urutan pertama peningkatan kecelakaan menurut data

Global Status Report on Road Safety yang dikeluarkan WHO. Indonesia salah satu negara yang memiliki jumlah penduduk terbesar di dunia, seiring dengan penambahan penduduk maka semakin tinggi pula tingkat mobilitas masyarakat. Mayoritas penduduk terpusat di pulau Jawa, dengan terpusatnya kegiatan dipulau Jawa maka meningkatkan mobilitas masyarakat yang menggunakan kendaraan. Indonesia dilaporkan mengalami kenaikan jumlah kecelakaan lalu lintas hingga lebih dari 80 persen. Menurut data Badan Pusat Statistik (BPS) pada tahun 2017 terjadi sebanyak 103.228 kejadian kecelakaan lalu lintas, dengan jumlah korban meninggal dunia sebanyak 30.568 jiwa, 14.395 mengalami luka berat, dan 119.945 mengalami luka ringan. Tiga provinsi di Jawa yaitu Jawa Barat, Jawa Tengah dan Jawa Timur memiliki tingkat kecelakaan tertinggi dibandingkan dengan provinsi lain di Indonesia. Jawa Tengah terdiri dari 35 kabupaten/kota dengan proyeksi penduduk tahun 2010-2020 menurut Badan Pusat Statistik (BPS) provinsi Jawa Tengah tahun 2016 sebesar 34.257.865 jiwa, data tersebut menggambarkan tingkat kepadatan penduduk di Jawa Tengah. Pada periode 1 Januari hingga 15 Oktober tahun 2017 Jawa Tengah menempati posisi di bawah Jawa Timur dengan jumlah kejadian kecelakaan sebanyak 14.087 kejadian dan 2.290 korban meninggal dunia. Menurut informasi dari NTMC KORLANTAS melalui program *Integrates Road Safety Management System* (IRSMS) 5 POLRES di POLDA Jawa Tengah melaporkan menggunakan aplikasi *Online IRSMS* dikategorikan berdasarkan tingkat fatalitas korban kecelakaan lalu lintas meninggal dunia yaitu RES Kabupaten Banyumas dengan 181 jiwa meninggal dunia, RES Cilacap 168 jiwa meninggal dunia,

RESTABES Semarang 162 jiwa meninggal dunia, RES Pati 153 jiwa meninggal dunia, dan RES Brebes 142 jiwa meninggal jiwa. Tetapi Kabupaten Pati menempati urutan pertama dalam jumlah kejadian kecelakaan, sebanyak 832 kejadian (NTMC KORLANTAS POLRI, 2017). Rendahnya kesadaran dalam keselamatan berkendara menjadi pemicu timbulnya kecelakaan lalu lintas. Banyak faktor penyebab dari kecelakaan lalu lintas, dari faktor pengguna, jalan, kendaraan dan lingkungan. Kasus kecelakaan lalu lintas dikawasan kabupaten Pati tahun 2017 cukup tinggi, data bersumber dari kantor SATLANTAS unit LAKA POLRES Pati menunjukkan bahwa banyaknya pelanggaran yang dilakukan oleh pengendara masih menjadi faktor terbesar penyebab kecelakaan seperti tidak patuhnya terhadap rambu lalu lintas, tidak menggunakan atribut berkendara sesuai peraturan, pengendara dibawah umur, kelelahan dalam berkendara, dan memacu kendaraan dengan kecepatan tinggi. Dalam periode 1 Januari hingga 29 Maret 2017 terjadi kasus kecelakaan sebanyak 250 kejadian yang menyebabkan 46 korban meninggal dunia. Korban meninggal dunia rata-rata masih dalam umur produktif yaitu 15-30 tahun (MURIANEWS.COM, 2017). Dengan banyaknya kejadian kecelakaan dan tingginya korban kecelakaan yang terjadi di Kabupaten Pati, maka harus ada tindak lanjut untuk menekan ataupun meminimalisir meningkatnya jumlah kejadian maupun korban. Berdasarkan uraian latar belakang diatas, peneliti tertarik untuk melakukan penelitian karya tulis yang berjudul “Model Regresi Logistik Biner *Bootstrap* Korban Kecelakaan Lalu Lintas di Kabupaten Pati”

1.2 Perumusan Masalah

Dari uraian latar belakang di atas, peneliti merumuskan beberapa masalah yang akan diteliti di dalam karya tulis ini yaitu sebagai berikut:

1. Bagaimana hasil permodelan regresi logistik biner kasus kecelakaan lalu lintas di kabupaten Pati?
2. Mencari nilai *standard error* berdasarkan model regresi logistik biner dengan menggunakan metode *Resampling Bootstrap*?

1.3 Tujuan Penelitian

Tujuan penelitian ini adalah ingin menjawab permasalahan dalam rumusan masalah. Adapun tujuan yang ingin dicapai oleh peneliti dalam penelitian ini adalah:

1. Memodelkan dengan regresi logistik biner kasus kecelakaan lalu lintas di Kabupaten Pati.
2. Mendapatkan nilai *standard error* terkecil berdasarkan model regresi logistik biner yang telah menggunakan metode *Resampling Bootstrap*.

1.4 Pembatasan Masalah

Agar permasalahan dalam penelitian fokus membahas kasus, maka diperlukan pembatasan masalah. Data logistik biner berbentuk kategorik dibahas dalam karya tulis ini lebih spesifik, maka peneliti membatasi pembahasan permasalahan. Dalam karya tulis ini, pembahasan akan difokuskan untuk membandingkan hasil analisis menggunakan regres logistk biner dengan regresi logistik biner yang di tambahkan

dengan *Resampling Bootstrap*. Bentuk data yang dapat diolah menggunakan *resampling Bootstrap* adalah variabel dependen (Y) berupa data dikotomus. Untuk memperoleh hasil output dengan mudah, analisis dilakukan dengan menggunakan bantuan *software R*. Data dependen yang digunakan dalam penelitian ini adalah keadaan korban kecelakaan lalu lintas dengan kategori, kategori 1 untuk korban kecelakaan hidup, kategori 0 untuk korban kecelakaan meninggal dunia. Sedangkan variabel independennya yaitu usia, waktu kejadian, jenis kecelakaan, jenis kendaraan, faktor manusia, atribut berkendara, dan peran korban.

1.5 Manfaat Penelitian

Manfaat penelitian ini dibagi menjadi empat yaitu:

a. Bagi Universitas Muhammadiyah Semarang

Menambah literatur mengenai analisis regresi logistik biner dengan *Resampling Bootstrap* dan dapat dipergunakan sebagai bahan bacaan dan referensi bagi pembaca dalam melakukan penelitian yang berhubungan dengan analisis regresi logistik biner dan *resampling Bootstrap*.

b. Bagi SATLANTAS unit LAKA POLRES Pati

Penelitian ini dapat memberikan saran atau sumbangsih pemikiran untuk Instansi sebagai penambah informasi.

c. Bagi Peneliti

Menambah pengalaman untuk melakukan penelitian dan mengamalkan keilmuan yang didapat dari proses belajar di perkuliahan.