

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dewasa ini permasalahan klasifikasi sering dijumpai di berbagai bidang dalam kehidupan sehari-hari diantaranya seperti dalam bidang ekonomi, perbankan, sosial, dan kesehatan. Mengingat pentingnya klasifikasi maka penelitian mengenai klasifikasi terus dikembangkan.

Pendekatan metode klasifikasi memiliki dua macam pendekatan, yaitu pendekatan parametrik dan nonparametrik. Metode dengan pendekatan parametrik sudah mulai ditinggalkan karena memiliki keterbatasan yaitu harus memenuhi beberapa asumsi seperti normalitas dan homogenitas varians. Penelitian metode klasifikasi dengan pendekatan nonparametrik lebih banyak dikembangkan mengingat metode tersebut memiliki fleksibilitas yang lebih besar dibandingkan dengan pendekatan parametrik karena tidak terikat pada asumsi tertentu. Metode klasifikasi nonparametrik salah satunya adalah metode klasifikasi dengan struktur pohon (Gordon,2013).

Metode pohon klasifikasi dan pohon regresi banyak dijumpai dalam berbagai penelitian diantaranya penelitian oleh Ture et al (2005) yang meneliti beberapa teknik pohon klasifikasi pada pengklasifikasian hipertensi di Klinik Kardiologi Fakultas Kedokteran Universitas Trakya di Turki, dan Handayani dan Purnami (2014) yang melakukan klasifikasi tingkat keganasan kanker tiroid dengan pendekatan

Classification and Regression Tree. Metode pohon klasifikasi dan regresi banyak digunakan karena diketahui mampu memberikan hasil prediksi dengan tingkat kesalahan yang kecil. Metode ini juga semakin populer karena memberikan kemudahan dalam interpretasi hasil analisisnya (Aeni,2009).

Metode pohon klasifikasi dan pohon regresi banyak digunakan dalam penelitian yang bersifat penggalian data (*data mining*). Dalam penggalian data terdapat beberapa teknik yang sering digunakan diantaranya *clustering*, *association rule*, *neural network*, dan pohon keputusan (*decision tree*). Pohon keputusan merupakan salah satu metode klasifikasi yang sering digunakan untuk penelitian dengan data yang memiliki banyak variabel independen. Algoritma yang sering digunakan dalam pohon keputusan diantaranya ID3, CART, dan C4.5 (Amri, 2013).S

Metode pohon klasifikasi yang saat ini sering digunakan adalah metode *Classification and Regression Tree* (CART). Kelebihan dari metode CART adalah metode ini tidak terikat oleh asumsi – asumsi dan efektif digunakan pada data yang memiliki dimensi besar atau memiliki banyak variabel independen. Akan tetapi metode CART memiliki kelemahan dimana metode tersebut kurang stabil pada perubahan data learning yang menyebabkan perubahan pada hasil prediksinya. Untuk mengatasi hal tersebut, dikembangkan suatu metode *ensemble* untuk meningkatkan akurasi prediksi (Timofeev,2004).

Metode *ensemble* merupakan gabungan beberapa dugaan dari beberapa pohon. Penelitian Sartono dan Syafitri (2010) menyatakan bahwa kelinieran hubungan antara peubah prediktor dan respon sering menjadi kendala dalam metode regresi klasik sehingga walaupun dilakukan sebuah transformasi hasilnya tetap tidak maksimal, oleh karena itu metode *ensemble* dapat dijadikan alternatif. Beberapa penelitian membuktikan bahwa keakurasian prediksi dapat ditingkatkan dengan metode *ensemble*, diantaranya penelitian oleh Wibowo dan Purwarianti (2011) yang melakukan metode *ensemble bagging* untuk memperbaiki hasil prediksi nasabah perusahaan asuransi X, dan penelitian Fernanda dan Otok (2012) yang melakukan *boosting* CART pada klasifikasi diabetes mellitus tipe II.

Metode *ensemble* mempunyai dua buah metode yang sering digunakan yaitu metode *bagging* dan *boosting*. Metode *bagging* diperkenalkan pada tahun 1996 oleh Breiman sedangkan *boosting* diperkenalkan oleh Freund dan Schapire pada tahun 1996. Menurut Breiman (1996), kelebihan dari metode *bagging* adalah mampu meningkatkan akurasi dengan mengurangi galat baku dari dugaan *classifier tunggal*, sedangkan kekurangannya adalah metode *bagging* tidak memberikan akurasi yang baik pada data dengan kategori peubah respon yang tidak seimbang. Menurut Liaw dan Wiener (2012), metode *boosting* mampu meningkatkan akurasi lebih baik dari metode *bagging* akan tetapi iterasi komputasi pada metode *bagging* lebih rumit dari metode *boosting*.

Kekurangan yang ditemukan pada metode *bagging* dan *boosting* menjadikan metode *ensemble* terus dikembangkan. Metode *ensemble* terbaru diciptakan oleh Breiman pada tahun 2003 yaitu metode *random forest*. Kelebihan dari metode *random forest* adalah proses iterasi komputasi yang lebih cepat, dapat digunakan pada data sampel kecil tapi berdimensi besar, dan pada data dengan kategori peubah respon yang tidak seimbang (Liaw dan Wiener, 2002).

Penelitian mengenai kemampuan akurasi masing – masing metode *ensemble* dibuktikan dalam penelitian oleh Muttaqien dkk (2013) yang membandingkan keakurasian beberapa metode *ensemble* dimana hasilnya jika hanya dilakukan satu buah metode *classifier* (CART) maka nilai akurasi yang didapatkan sangat rendah yaitu sebesar 5,02% sedangkan bila dilakukan suatu metode *ensemble* terjadi peningkatan akurasi yang jauh lebih tinggi yaitu : 19,26% untuk metode *bagging*, 75,89% untuk metode *boosting*, dan 94,66% untuk *random forest* .

Akurasi yang tinggi pada bidang kesehatan sangatlah penting karena berhubungan erat dengan kelangsungan hidup manusia. Penelitian yang akurat dapat membantu pemerintah dalam pengambilan keputusan yang efektif sehingga kesejahteraan umum dapat tercapai. Salah satu unsur dari kesejahteraan umum adalah terwujudnya derajat kesehatan masyarakat yang setinggi-tingginya termasuk kesehatan gigi. Gigi merupakan salah satu organ tubuh manusia yang memegang peranan penting dalam fungsi pencernaan. Oleh sebab itu kesehatan gigi harus dijaga mulai sedini mungkin (Tampubolon,2005).

Survei Kesehatan Rumah Tangga (SKRT) 2012 memnyatakan bahwa angka penyakit gigi dan mulut yang diderita masyarakat Indonesia mencapai 90,05%, artinya hampir seluruh masyarakat Indonesia pernah mengalami penyakit gigi dan mulut. Ada berbagai macam masalah pada gigi dan mulut yang sering dijumpai pada masyarakat salah satunya adalah penyakit jaringan penyangga gigi atau penyakit periodontal. Penyakit periodontal umumnya disebabkan oleh kebersihan mulut yang buruk. Akumulasi plak yang berisi mikroorganisme pada gigi dapat menyebabkan akumulasi karang gigi yang dapat mengganggu jaringan penyangga gigi atau gusi. Radang pada jaringan penyangga gigi disebut periodontitis (Kun dan Kusuma, 2013).

Penyakit periodontitis dan karies gigi adalah penyakit gigi terbanyak yang diderita masyarakat Indonesia sehingga memerlukan perhatian serius. Umumnya penyakit periodontal dikeluhkan masyarakat dalam bentuk gusi berdarah, gigi goyang, gusi bengkak, dan bau mulut (Sudiby,2003). Menurut Departemen Kesehatan Republik Indonesia (1999), berdasarkan hasil survey penyakit periodontal tahun 1994-1999 menunjukkan bahwa penyakit periodontal menyerang 80,8% anak Indonesia di bawah usia 12 tahun.

Tingginya prevalensi penyakit periodontitis mengakibatkan upaya pencegahan dan penanganan terus dikembangkan. Pemerintah melalui Puskesmas telah melaksanakan berbagai kegiatan dalam rangka meningkatkan derajat kesehatan gigi dan mulut, diantaranya adalah melaksanakan upaya pelayanan kesehatan gigi pencegahan.

Upaya kesehatan gigi puskesmas sampai saat ini belum dapat berjalan dengan optimal oleh karena adanya berbagai kendala, antara lain : keterbatasan tenaga, sarana, biaya operasional maupun kondisi sosial dan ekonomi masyarakat (Depkes RI,2000). Upaya pencegahan penyakit periodontitis memerlukan strategi yang komprehensif mengingat penyakit periodontitis dapat terjadi akibat beberapa faktor diantaranya keadaan rongga mulut seperti akumulasi plak dan karang gigi juga perilaku kesehatan seperti kebiasaan menyikat gigi dan kebiasaan merokok. Penelitian yang memiliki akurasi tinggi mengenai faktor yang mempengaruhi terjadinya periodontitis kronis sesuai dengan klasifikasi ringan, sedang, dan berat dapat dilakukan untuk membantu dalam pengambilan keputusan yang efektif. Perlu dilakukan kajian mengenai penggunaan Random Forest dan Classification and Regression Tree (RF-CART) Untuk Klasifikasi Perodontitis Kronis.

1.2 Rumusan Masalah

Permasalahan dalam penelitian, yaitu:

1. Bagaimana keakurasian penerapan metode CART pada klasifikasi perodontitis Kronis Pasien Rumah Sakit Gigi dan Mulut Universitas Hang Tuah Surabaya?
2. Bagaimana keakurasian penerapan metode Random Forest dan CART (RF-CART) pada klasifikasi klasifikasi perodontitis Kronis Pasien Rumah Sakit Gigi dan Mulut Universitas Hang Tuah Surabaya?

3. Bagaimana peningkatan keakurasian penggunaan metode RF-CART dibandingkan dengan penggunaan metode CART pada klasifikasi periodontitis Kronis Pasien Rumah Sakit Gigi dan Mulut Universitas Hang Tuah Surabaya?

1.3 Tujuan Penelitian

1.3.1 Tujuan Umum

Melakukan klasifikasi periodontitis kronis dan mengidentifikasi faktor-faktor yang mempengaruhi periodontitis kronis pada pasien Rumah Sakit Gigi dan Mulut Universitas Hang Tuah Surabaya menggunakan metode RF-CART.

1.3.2 Tujuan Khusus

1. Menerapkan metode CART pada klasifikasi periodontitis Kronis Pasien Rumah Sakit Gigi dan Mulut Universitas Hang Tuah Surabaya dan menghitung nilai akurasinya.
2. Menerapkan metode Random Forest dan CART (RF-CART) pada klasifikasi periodontitis Kronis Pasien Rumah Sakit Gigi dan Mulut Universitas Hang Tuah Surabaya dan menghitung nilai akurasinya.
3. Membandingkan keakurasian metode RF-CART dengan metode CART pada klasifikasi periodontitis Kronis Pasien Rumah Sakit Gigi dan Mulut Universitas Hang Tuah Surabaya.

1.4 Manfaat Penelitian

1.4.1 Bagi Penyedia Kebijakan dan Layanan Kesehatan Gigi dan Mulut

Hasil penelitian dapat digunakan sebagai pertimbangan penyusunan kebijakan kesehatan yang efektif sebagai upaya pencegahan penyakit periodontal.

1.4.2 Bagi Perkembangan Ilmu Kesehatan Gigi Masyarakat

Hasil penelitian dapat menjadi pengembangan Ilmu Kesehatan Gigi Masyarakat yang dalam upaya preventif yang terkait kesehatan gigi pada masyarakat.

1.4.2 Bagi Penelitian Selanjutnya

Hasil penelitian diharapkan dapat menjadi masukan dan rujukan bagi penelitian selanjutnya yang sejenis.